

# MARKOV RANDOM FIELDS, MARKOV COCYCLES AND THE 3-COLORED CHESSBOARD

NISHANT CHANDGOTIA AND TOM MEYEROVITCH

**ABSTRACT.** The well-known Hammersley-Clifford theorem states (under certain conditions) that any Markov random field is a Gibbs state for a nearest neighbor interaction. In this paper we study Markov random fields for which the proof of the Hammersley-Clifford theorem does not apply. Following Petersen and Schmidt we utilize the formalism of cocycles for the homoclinic equivalence relation and introduce “Markov cocycles”, reparametrisations of Markov specifications. The main part of this paper exploits this to deduce the conclusion of the Hammersley-Clifford theorem for a family of Markov fields which are outside the theorem’s purview where the underlying graph is  $\mathbb{Z}^d$ . This family includes all Markov random fields whose support is the  $d$ -dimensional “3-colored chessboard”. On the other extreme, we construct a family of shift-invariant Markov random fields which are not given by any finite range shift-invariant interaction.

## 1. INTRODUCTION

A Markov random field is a collection of random variables  $(x_v)_{v \in V}$  indexed by the vertices of an undirected graph  $G = (V, E)$ , satisfying the following conditional independence property: Any two subsets of the variables  $\{x_v\}_{v \in A}$  and  $\{x_v\}_{v \in B}$  corresponding to separated subsets of vertices  $A, B \subset V$  in the graph are independent given  $\{x_v\}_{v \in V \setminus (A \cup B)}$ . Under a certain positivity assumption on the joint distribution, the Hammersley-Clifford theorem states that any Markov random field is a Gibbs state for a nearest neighbor interaction. This means that the distribution of configurations on any finite subset  $A \subset V$  given the configuration on the outer boundary can be expressed as the normalized product of weights associated with configurations on complete subgraphs. Moreover, if the measure is invariant under a group of automorphisms of the graph, the weights can be chosen invariant under the same. The key assumption in the proof of the Hammersley-Clifford theorem is what we call existence of a “safe symbol”.

In this paper, we shall study Markov random fields outside the purview of the safe symbol assumption. We focus on Markov random fields which consist of random variables taking values in a finite set  $\mathcal{A}$ . Thus, a Markov random field on the graph  $G = (V, E)$  is a probability measure  $\mu$  on  $\mathcal{A}^V$ , satisfying the conditional independence property described above. The *support* of  $\mu$  is the smallest closed subset  $X \subset \mathcal{A}^V$  such that  $\mu(X) = 1$ . The support  $X$  of a Markov random field has a certain property which makes it a “topological Markov field” (see section 2.1). A specification is a consistent system of probability distributions on configurations for every finite set of sites given the configuration on the complement. A Markov specification is a specification where the distribution depends solely on the configuration on the outer boundary of the concerned finite set and Gibbs if it can be written as the normalized product of nearest neighbor interactions as well. The Hammersley-Clifford theorem is thus a statement about the specification of a Markov random field, under an assumption on its support (existence of a “safe symbol”).

---

2010 *Mathematics Subject Classification.* Primary 37A60; Secondary 60J99.

*Key words and phrases.* Markov random field, Gibbs state, nearest neighbour interaction, shift-invariance, multidimensional, cocycle, pivot property.

In general the conclusion of the Hammersley-Clifford theorem does not hold if we drop the safe symbol assumption: J. Moussouris provided examples of Markov random fields on a finite graph which are not Gibbs states for a nearest neighbor interaction [16]. Are there other conditions on a Markov random field which guarantee that a Hammersley-Clifford type theorem to holds? In the absence of a nearest neighbor interaction, is there an alternative “compact description” for the specification of a Markov random field? These are the types of questions we investigate in this paper. Some work along these lines has been carried out in the first-named author’s M.Sc. thesis [5].

There has been other work along this direction: When the underlying graph is finite, there are known algebraic conditions on the support [12] and conditions on the graph [15] for the conclusions of the Hammersley-Clifford theorem to hold. When the underlying graph is a Cayley graph of  $\mathbb{Z}$  the Markov random field is shift-invariant and  $\mathcal{A}$  is finite, the conclusions of the Hammersley-Clifford theorem holds without any further assumptions [6]. Furthermore, in that setting any Markov random field is a stationary Markov chain. Even when the underlying graph is a Cayley graph of  $\mathbb{Z}$ , this conclusion can fail for countable  $\mathcal{A}$  (Theorem 11.33[11]), or if we drop the assumption of shift-invariance [9]. In the same setting, except  $\mathcal{A}$  is a general measure space rather than a finite set, certain mixing conditions guarantee the conclusion (theorems 10.25 and 10.35 in [11]). When the underlying graph is the Cayley graph of  $\mathbb{Z}^d$  and  $d > 1$ , the conclusion fails even in the shift-invariant and finite alphabet case (Chapter 5 in [5] and section 9 of this paper).

For most of this paper we focus on Markov random fields for which the graph  $G = (V, E)$  is the Cayley graph of  $\mathbb{Z}^d$  with respect to the standard set of generators. The natural action of  $\mathbb{Z}^d$  on this graph induces an action on  $\mathcal{A}^{\mathbb{Z}^d}$  by translations or *shifts*. We study in particular Markov random fields which are invariant to the shift action. Evidently, the support of any such Markov random field is a *shift space*, that is a shift invariant compact subset of  $\mathcal{A}^{\mathbb{Z}^d}$ .

A major portion of this paper is spent establishing the conclusion of the Hammersley-Clifford theorem for Markov random fields whose support is in a certain one-parameter family of shift spaces  $X_r$ , where  $r \geq 2$  is an integer. We prove that any Markov random field whose support is  $X_r$  is a Gibbs state for a nearest neighbor interaction (proposition 5.5). Furthermore, the interaction can be chosen shift-invariant if the concerned Markov random field is shift-invariant (theorem 6.1). For  $r \neq 4$ , the space  $X_r \subset \{0, \dots, r-1\}^{\mathbb{Z}^d}$  consists of all those  $x \in \{0, \dots, r-1\}^{\mathbb{Z}^d}$  for which  $x_n - x_m = \pm 1 \pmod r$  whenever  $\|n - m\|_1 = 1$ . For  $r = 2$  this space consists of 2 periodic “chessboard” configurations and for  $r = 3$  this is the  $d$ -dimensional “3-colored chessboard”: the set of proper graph 3-colorings of the Cayley graph of  $\mathbb{Z}^d$ .

In section 3 we introduce “Markov cocycles” and “Gibbs cocycles”, a rather general formalism allowing to parameterize Markov specifications. Following [18], these are logarithms of the Radon-Nikodym derivatives with respect to the homoclinic relation, that is, logarithm of the ratio of probabilities of configurations which differ at only finitely many sites. Markov cocycles form a vector space, the space of Gibbs cocycles with shift-invariant interactions being finite dimensional. A strong version of Hammersley-Clifford theorem can now be stated in terms of these objects (theorem 3.1). A weaker condition on the support of a Markov random field than that of a “safe symbol”, namely, the “pivot property” gives us a weaker result, that is, the space of shift-invariant Markov cocycles is finite dimensional. For the spaces  $X_r$ ’s we show that the dimension of the shift-invariant Markov cocycles is  $r$

while the dimension of Gibbs cocycles with a shift-invariant interaction is  $r - 1$  (propositions 5.2 and 5.3). However if we do not demand shift-invariance for the interaction, then every Markov cocycle is Gibbs. Thus for shift spaces  $X_r$ 's the conclusion of the weaker version of Hammersley-Clifford theorem holds, the conclusion of the stronger version holds as well but not in the shift-invariance setting. We give an example in Section 9 which does not have the pivot property and the space of shift-invariant Markov cocycles has uncountable dimension.

## 2. BACKGROUND AND NOTATION

This section will recall the necessary concepts and introduce the basic notation.

**2.1. Markov Random Fields and topological Markov fields.** Let  $G = (V, E)$  be a graph where the vertex set  $V$  is finite or countable. The *boundary* of a set of vertices  $F \subset V$ , denoted  $\partial F$ , is the set of vertices outside  $F$  which are adjacent to  $F$ :

$$\partial F := \{v \in V \setminus F : \exists w \in F \text{ s.t. } (v, w) \in E\}.$$

**Remark:** Observe that in our notation  $\partial F \subset F^c$ . This is sometimes called the *outer boundary* of a set  $F$ . Consistent with our notation the *inner boundary* of  $F$  is  $\partial F^c$ .

Given  $\mathcal{A}$  finite (more generally compact),  $\mathcal{A}^V$  is a compact topological space with respect to the product topology. For  $F \subset V$  finite and  $a \in \mathcal{A}^F$ , we denote by  $[a]_F$  the *cylinder set*

$$[a]_F = \{x \in \mathcal{A}^V : x|_F = a\}.$$

The collection of cylinder sets generate the Borel  $\sigma$ -algebra on  $\mathcal{A}^V$ .

A *Markov random field* is a Borel probability measure  $\mu$  on  $\mathcal{A}^V$  with the property that for all finite  $A, B \subset V$  such that  $\partial A \subset B \subset A^c$  and  $a \in \mathcal{A}^A, b \in \mathcal{A}^B$  satisfying  $\mu([b]_B) > 0$

$$\mu([a]_A \mid [b]_B) = \mu([a]_A \mid [b]_{\partial A}).$$

An equivalent definition is the following: If  $x$  is a point chosen randomly according to the measure  $\mu$ , and  $A, B \subset V$  are finite, *separated* sets in  $G$  (meaning  $(v, w) \notin E$  whenever  $v \in A$  and  $w \in B$ ), then conditioned on  $x|_{V \setminus (A \cup B)}$ ,  $x|_A$  and  $x|_B$  are independent random variables.

**Remark:** A stronger notion of a Markov random field obtained by requiring the conditional independence above for all  $A, B \subset V$  which are separated in  $G$ , finite or not is called a *global Markov random field*. We will not discuss this distinction further here, and be mostly concerned with the definition of Markov random fields where the above mentioned  $A$  and  $B$  are finite.

As in [5, 6], a *topological Markov field* is a compact set  $X \subset \mathcal{A}^V$  such that for all  $F \subset V$  finite, and all  $x, y \in X$  satisfying  $x = y$  on  $\partial F$ ,  $z \in \mathcal{A}^V$  given by

$$z_v = \begin{cases} x_v & \text{for } v \in F \\ y_v & \text{for } v \in V \setminus F \end{cases}$$

is an element in  $X$ . A topological Markov field is called *global* if we do not demand that  $F$  be finite.

The *support* of a Borel probability measure  $\mu$  on  $\mathcal{A}^V$  denoted by  $\text{supp}(\mu)$  is the intersection of all closed sets  $Y \subset \mathcal{A}^V$  for which  $\mu(Y) = 1$ . Equivalently,

$$\text{supp}(\mu) = \mathcal{A}^V \setminus \bigcup_{[a]_A \in \mathcal{N}(\mu)} [a]_A,$$

where  $\mathcal{N}(\mu)$  is the collection of all cylinder sets with  $\mu([a]_A) = 0$ . It follows that the support of a Markov random field is a topological Markov field.

**2.2. The homoclinic equivalence relation of a TMF and adapted MRFs.** Following [18, 21], we denote by  $\Delta_X$  the *homoclinic equivalence relation* of a TMF  $X \subset \mathcal{A}^V$ , which is given by

$$\Delta_X := \{(x, y) \in X \times X \mid x_n = y_n \text{ for all but finitely many } n \in V\}.$$

We say that a Markov random field  $\mu$  is *adapted* with respect to topological Markov field  $X$  if  $\text{supp}(\mu) \subset X$  and

$$x \in \text{supp}(\mu) \implies \{y \in X \mid (x, y) \in \Delta_X\} \subset \text{supp}(\mu).$$

Equivalently, a Markov random field  $\mu$  is adapted to  $X$  if and only if the measure  $\mu$  is non-singular with respect to  $\Delta_X$ .

To illustrate this definition, if  $\text{supp}(\mu) = X$  then  $\mu$  is adapted with respect to  $X$ , and if  $X = X_1 \cup X_2$  is the union of topological Markov fields over disjoint alphabets and  $\mu$  is a Markov random field with  $\text{supp}(\mu) = X_1$  then  $\mu$  is adapted with respect to  $X$ . On the other hand, the Bernoulli measure  $(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1)^V$  is *not* adapted with respect to  $\{0, 1, 2\}^V$ . In fact, if  $X = \mathcal{A}^{\mathbb{Z}^d}$  then any Markov random field which is adapted to  $X$  has  $\text{supp}(\mu) = X$ .

**2.3.  $\mathbb{Z}^d$ -shift spaces and shifts of finite type.** For the Markov random fields we discuss in this paper, the vertices of the underlying graph are the  $d$ -dimensional integer lattice. We identify  $\mathbb{Z}^d$  with the set of vertices of the Cayley graph with respect to the standard generators. Rephrasing,  $n, m \in \mathbb{Z}^d$  are adjacent iff  $\|n - m\|_1 = 1$ , where for any  $n = (n_1, \dots, n_r) \in \mathbb{Z}^d$ ,  $\|n\|_1 = \sum_{r=1}^d |n_r|$  denotes the  $l_1$  norm of  $n$ . The *boundary* of any finite set  $F \subset \mathbb{Z}^d$  is thus given by:

$$\partial F = \{m \in F^c \mid \|n - m\|_1 = 1 \text{ for some } n \in F\}.$$

On the compact space  $\mathcal{A}^{\mathbb{Z}^d}$  the maps  $\sigma_n : \mathcal{A}^{\mathbb{Z}^d} \longrightarrow \mathcal{A}^{\mathbb{Z}^d}$  given by

$$(\sigma_n(x))_m = x_{m+n} \text{ for all } m, n \in \mathbb{Z}^d$$

define a  $\mathbb{Z}^d$ -action by homeomorphisms, called the *shift-action*. The pair  $(\mathcal{A}^{\mathbb{Z}^d}, \sigma)$  is a topological dynamical system called the *d-dimensional full shift* on the alphabet  $\mathcal{A}$ . Note that  $\sigma$  acts on the Cayley graph of  $\mathbb{Z}^d$  by graph isomorphisms.

A  $\mathbb{Z}^d$  *shift space* or *subshift* is a dynamical system  $(X, \sigma)$  where  $X \subset \mathcal{A}^{\mathbb{Z}^d}$  is closed and invariant under the map  $\sigma_n$  for each  $n \in \mathbb{Z}^d$ .

The *language* of a subshift  $X$  denoted by  $\mathcal{B}(X)$  is defined as all finite patterns which occur in the elements of  $X$ . For  $W \subset \mathbb{Z}^d$  let

$$\mathcal{B}_W(X) = \{w \in \mathcal{A}^W \mid \text{there exists } x \in X \text{ such that } x|_W = w\}.$$

Therefore

$$\mathcal{B}(X) = \cup_{W \subset \mathbb{Z}^d \text{ finite}} \mathcal{B}_W(X).$$

If  $A, B \subset \mathbb{Z}^d$  are disjoint sets,  $x \in \mathcal{A}^A, y \in \mathcal{A}^B$ , then  $x \vee y \in \mathcal{A}^{A \cup B}$  is given by

$$x \vee y = \begin{cases} x_n & n \in A \\ y_n & n \in B. \end{cases}$$

A Borel probability measure  $\mu$  on  $\mathcal{A}^{\mathbb{Z}^d}$  is *shift-invariant* if  $\mu \circ \sigma_n = \mu$  for all  $n \in \mathbb{Z}^d$ . It follows that the support of any shift-invariant measure  $\mu$  is a subshift.

An alternate equivalent definition for a subshift is given by *forbidden patterns* as follows: Let

$$\mathcal{A}^* = \{\mathcal{A}^W \mid W \text{ is a finite subset of } \mathbb{Z}^d\}.$$

For any  $\mathcal{F} \subset \mathcal{A}^*$  let

$$X_{\mathcal{F}} = \{x \in \mathcal{A}^{\mathbb{Z}^d} \mid \text{no translate of a subword of } x \text{ belongs to } \mathcal{F}\}.$$

Evidently, a subset  $X \subset \mathcal{A}^{\mathbb{Z}^d}$  is a subshift if and only if there exists  $\mathcal{F} \subset \mathcal{A}^*$  such that  $X = X_{\mathcal{F}}$ . The set  $\mathcal{F}$  is called the set of forbidden patterns for  $X$ . A subshift  $X$  is called a *shift of finite type* if  $X = X_{\mathcal{F}}$  for some finite set  $\mathcal{F}$ . A shift of finite type is called a *nearest neighbor shift of finite type* if  $X = X_{\mathcal{F}}$  where  $\mathcal{F}$  consists of nearest neighbor constraints, i.e.  $\mathcal{F}$  consists of patterns on edges. When  $d = 1$  nearest neighbor shift of finite type are also called *topological Markov chains*. In fact the study of nearest neighbor shifts of finite type has been much motivated by the fact that the support of stationary Markov chains are one-dimensional nearest neighbor shifts of finite type.

Every nearest neighbor  $\mathbb{Z}^d$ -shift of finite type is a shift-invariant topological Markov field. When  $d = 1$  the converse is also true under the assumption that the subshift is non-wandering [6]. Without the non-wandering assumption, one-dimensional shift-invariant topological Markov fields are sofic shifts, but not necessarily of finite type [6]. This does not hold in higher dimensions ([5] and section 9).

**2.4. Gibbs states with nearest neighbor interactions.** For a graph  $G = (V, E)$  and any set  $A \subset V$ , let  $\text{diam}(A)$  denote the diameter of  $A$  with respect to the graph distance in  $G$ , that is,  $\text{diam}(A) = \max_{i,j \in A} \|i - j\|$ .

Following [19], an *interaction* on  $X$  is a function  $\phi$  from  $\mathcal{B}(X)$  to  $\mathbb{R}$ , satisfying certain summability conditions. Here we will only consider finite range interactions, for which the summability conditions are automatically satisfied.

An interaction is of *range at most  $k$*  if  $\phi(a) = 0$  for  $a \in \mathcal{B}_A(X)$  whenever  $\text{diam}(A) \geq k$ . We will call an interaction of range 1 a *nearest neighbor interaction*. When  $G = \mathbb{Z}^d$ , an interaction  $\phi$  is shift-invariant if for all  $n \in \mathbb{Z}^d$  and  $a \in \mathcal{B}(X)$ ,  $V(a) = V(\sigma_n(a))$ . Since the standard Cayley graph of  $\mathbb{Z}^d$  has no triangles, a shift-invariant nearest neighbor interaction is uniquely determined by its values on configurations on  $\{0\}$  (“single site potentials”) and on configurations on pairs  $\{0, e_i\}$ ,  $i = 1, \dots, d$  (“edge interactions”). We denote these by  $\phi([a]_0)$  and  $\phi([a, b]_i)$  respectively where  $a, b \in \mathcal{A}$ .

A *Gibbs state with a nearest neighbor interaction*  $\phi$  is a Markov random field  $\mu$  such that for all  $x \in \text{supp}(\mu)$  and  $A, B \subset \mathbb{Z}^d$  finite satisfying  $\partial A \subset B \subset A^c$

$$\mu([x]_A \mid [x]_B) = \frac{\prod_{C \subset A \cup \partial A} e^{\phi([x]_C)}}{Z_{A,x|\partial A}}$$

where  $Z_{A,x|\partial A}$  is the uniquely determined normalizing factor so that  $\mu(X) = 1$ , dependent upon  $A$  and  $x|_{\partial A}$ .

### 3. MARKOV SPECIFICATIONS AND MARKOV COCYCLES

Any Markov random field  $\mu$  yields conditional probabilities of the form  $\mu(x_F = \cdot \mid x_{\partial F} = \delta)$  for any finite  $F \subset V$  and all *admissible*  $\delta \in \mathcal{A}^{\partial F}$  (by admissible we mean  $\mu([\delta]_{\partial F}) > 0$ ). We refer to such collection of conditional probabilities as the *Markov specification* associated with  $\mu$ . It may happen that two distinct Markov random fields have the same specification, as in the case of the 2-dimensional Ising model in low temperature [17]. In general it is a subtle and challenging problem to determine if a given Markov specification admits more than one Markov random fields (the problem of uniqueness for the measure of maximal entropy of a  $\mathbb{Z}^d$ -shift of finite type is an instance of this problem [4]). For the purpose of our study and statement of our results, it would be convenient to have an intrinsic definition for a Markov specification, not involving a particular underlying Markov random field.

Let  $X \subset \mathcal{A}^{\mathbb{Z}^d}$  be a topological Markov field. A *Markov specification* on  $X$  is an assignment for each finite  $F \subset \mathbb{Z}^d$  and  $x \in \mathcal{B}_{\partial F}(X)$  of a probability measure  $\Theta_{F,x}$  on  $\mathcal{B}_F(X)$  satisfying the following conditions:

- (1) **Support condition:** For any finite  $F \subset V$ , any  $x \in \mathcal{B}_{\partial F}(X)$  and any  $y \in \mathcal{B}_F(X)$  such that  $x \vee y \in \mathcal{B}_{F \cup \partial F}(X)$ ,  $\Theta_{F,x}(y) > 0$ . This condition can be written as follows:

$$\text{supp}(\Theta_{F,x}) = \{y \in \mathcal{B}_F(X) : x \vee y \in \mathcal{B}_{F \cup \partial F}(X)\}.$$

- (2) **Markovian condition:** For any finite  $F \subset V$  and  $x \in \mathcal{B}_{\partial F}(X)$ ,  $\Theta_{F,x}$  is a Markov random field on the finite graph induced from  $V$  on  $F$ .
- (3) **Consistency condition:** If  $F \subset H$ ,  $x \in \mathcal{B}_{\partial F}(X)$ ,  $y \in \mathcal{B}_{\partial H}(X)$  and  $x_n = y_n$  for  $n \in \partial F \cap \partial H$ , then for any  $z \in \mathcal{B}_F(X)$

$$\Theta_{F,x}(z) = \frac{\Theta_{H,y}([z \vee x]_{F \cup \partial F})}{\Theta_{H,y}([x]_{\partial F})}.$$

The definition above have been set up so that for any Markov random field  $\mu$  with  $X = \text{supp}(\mu)$ , the conditional probabilities on  $\mathcal{B}_F(X)$  obtained by conditioning on  $\partial F$  is indeed a Markov specification. Conversely, given any Markov specification  $\Theta$  on  $X$  there exists a Markov random field  $\mu$  on  $X$  compatible with  $\Theta$  in the sense that  $\mu([a]_F \mid [y]_{\partial F}) = \Theta_{F,y}(a)$ , whenever  $\mu([y]_{\partial F}) > 0$  [11]. Furthermore, for a specification  $\Theta$  which is shift-invariant, it follows from amenability of  $\mathbb{Z}^d$  that there exists a shift-invariant Markov random field  $\mu$  compatible with  $\Theta$ . However, in general it is possible that for a given specification  $\Theta$  the support of any  $\mu$  satisfying the above is contained in a strict subset of  $X$ , in which case there exist certain finite  $F \subset V$  and configurations  $y \in \mathcal{B}_{\partial F}(X)$  for which the conditional probabilities  $\mu([x]_F \mid [y]_{\partial F})$  are meaningless. We will provide such examples in Section 6. In such a case, according to our definition, the specification of  $\mu$  would be the restriction of  $\Theta$  to the support of  $\mu$ .

It will be convenient for our purposes to re-parameterize the set of Markov specifications on a given topological Markov field  $X$ . For this purpose we use the formalism of  $\Delta_X$ -cocycles. To this well-known formalism we introduce an ad-hoc definition of a *Markov cocycle*, which we now explain. As in [21], a (real-valued)  $\Delta_X$ -cocycle is a function  $M : \Delta_X \rightarrow \mathbb{R}$  satisfying

$$(3.1) \quad M(x, z) = M(x, y) + M(y, z) \text{ whenever } (x, y), (y, z) \in \Delta_X.$$

A  $\Delta_X$ -cocycle is shift-invariant if  $M(x, y) = M(\sigma_n(x), \sigma_n(y))$  for all  $n \in \mathbb{Z}^d$ . We call  $M$  a *Markov cocycle* if it is a  $\Delta_X$ -cocycle and satisfies: For any  $(x, y) \in \Delta_X$ ,  $M(x, y)$  depends

only upon the sites on which  $x$  and  $y$  differ and its boundary i.e. if  $F$  is the set of sites on which  $x$  and  $y$  differ then

$$M(x, y) = c_{(x|_{F \cup \partial F}, y|_{F \cup \partial F})}.$$

There is a clear bijection between Markov cocycles and Markov specifications on  $X$ : If  $\Theta$  is a Markov specification on  $X$ , the corresponding Markov cocycle is given by

$$M(x, y) = \log (\Theta_{F, y|_{\partial F}}(y|_F)) - \log (\Theta_{F, x|_{\partial F}}(x|_F)),$$

where  $(x, y) \in \Delta_X$  and  $F \subset \mathbb{Z}^d$  finite such that  $x_n = y_n$  for  $n \in F^c$ .

Conversely, given a Markov cocycle  $M$  on  $X$ , the corresponding specification  $\Theta$  is given by

$$\Theta_{F, a}(y) = \frac{1}{Z_{M, F, a}} e^{M(x \vee z, x \vee y)},$$

where  $F \subset \mathbb{Z}^d$  is a finite set,  $a \in \mathcal{B}_{\partial F}(X)$ ,  $y, z \in \mathcal{B}_F(x)$  are such that  $a \vee x, a \vee z \in \mathcal{B}_{F \cup \partial F}(X)$  and  $x \in \mathcal{B}_{F^c}(X)$  with  $x|_{\partial F} = a$ . The normalizing constant  $Z_{M, F, a}$  is given by:

$$Z_{M, F, a} = \sum_{y \in \mathcal{B}_{F \cup \partial F}(X) : y|_{\partial F} = a} e^{M(x \vee z, x \vee y)}.$$

Note that the expression for the specification is well defined since  $X$  is a topological Markov field.

This bijection clearly maps shift-invariant specifications to shift-invariant Markov-cocycles. Thus, a shift-invariant Borel probability measure  $\mu$  is a shift-invariant Markov random field if and only if  $X = \text{supp}(\mu)$  is a topological Markov field and there exists a shift-invariant Markov cocycle  $M$  on  $X$  such that for all  $(x, y) \in \Delta_X$

$$\frac{\mu([y]_\Lambda)}{\mu([x]_\Lambda)} = e^{M(x, y)}$$

for any  $\Lambda \supset F \cup \partial F$  where  $F$  is the set of sites on which  $x, y$  differ.

In the above situation the function  $\rho : \Delta_X \rightarrow \mathbb{R}_+$  given by  $\rho(x, y) = e^{M(x, y)}$  is the  $\Delta_X$ -Radon-Nikodym cocycle of  $\mu$  as in [18].  $\rho$  is clearly a cocycle taking values in the multiplicative group  $\mathbb{R}_+$ .

Given a topological Markov field  $X$ , the *Gibbs cocycle* corresponding to a nearest neighbor interaction  $\phi$  is the  $\Delta_X$ -cocycle given by:

$$M(x, y) = \sum_{W \subset V \text{ finite}} \phi(y|_W) - \phi(x|_W).$$

Note that  $M$  is well defined since there are only finitely many non-zero terms in the sum whenever  $(x, y) \in \Delta_X$ . Evidently, any  $M$  of this kind is a Markov cocycle. Our point of interest is the converse. Note that a Borel probability measure  $\mu$  is a Gibbs state with nearest neighbor interaction if and only if its  $\Delta_X$ -Radon-Nikodym cocycle is  $e^M$  for the Gibbs cocycle  $M$  on  $X = \text{supp}(\mu)$ .

Let  $X$  be a topological Markov field. We shall denote by  $\mathcal{M}_X$  the set of all Markov cocycles and by  $\mathcal{G}_X$  the set of all nearest neighbor Gibbs cocycles. In case  $X$  is also a subshift, we denote by  $\mathcal{M}_X^\sigma$  the set of shift-invariant Markov cocycles and by  $\mathcal{G}_X^\sigma$  the set of Gibbs cocycles corresponding to a shift-invariant nearest neighbor interaction.

The set  $\mathcal{M}_X$  of Markov cocycles naturally carries a vector space structure since given  $M_1, M_2 \in \mathcal{M}_X$  and  $c_1, c_2 \in \mathbb{R}$ ,

$$c_1 M_1 + c_2 M_2 \in \mathcal{M}_X.$$

It can be seen that  $\mathcal{G}_X$  is a subspace of  $\mathcal{M}_X$ . Since shift-invariant nearest neighbor interactions constitute a finite-dimensional vector space  $\mathcal{G}_X^\sigma \subset \mathcal{M}_X$  has finite dimension. We would like to note here that for finite graphs the condition under which a given Markov cocycle is Gibbs is very similar to the ‘balanced’ conditions mentioned in [16].

**3.1. The “safe symbol property” and the Hammersley-Clifford Theorem.** A subshift  $X$  on alphabet  $\mathcal{A}$  is said to have a *safe symbol* if there exists an element  $\star \in \mathcal{A}$  such that for all  $x \in X$  and  $A \subset \mathbb{Z}^d$ ,  $y \in \mathcal{A}^{\mathbb{Z}^d}$  given by

$$y_n = \begin{cases} x_n & \text{for } n \in A \\ \star & \text{for } n \in A^c. \end{cases}$$

is also an element of  $X$ .

A formulation of the Hammersley-Clifford Theorem states:

**Theorem. (Hammersley-Clifford, weak version [13])** *Let  $X$  be a topological Markov field with a safe symbol. Then:*

- (1) *Any Markov random field with  $\text{supp}(\mu) = X$  is a Gibbs state for a nearest neighbor interaction.*
- (2) *Any shift-invariant Markov random field on  $X$  is a Gibbs state for a shift-invariant, nearest neighbor interaction.*

The second statement in the theorem above is not a part of the original formulation, but does follow since there is an explicit algorithm to produce a nearest neighbor interaction which is invariant under any graph automorphism for which the original Markov random field was invariant [5]. See also [1],[23],[24]. It is in general false that a shift-invariant Markov random field whose (shift-invariant) specification is compatible with some nearest neighbor interaction must also be compatible with some shift-invariant nearest neighbor interaction (see corollary 5.6 below). In particular, for a general topological Markov field  $X$  we have  $\mathcal{G}_X^\sigma \subset \mathcal{M}_X^\sigma \cap \mathcal{G}_X$ , but the inclusion may be strict.

An inspection of the original proof of the Hammersley-Clifford theorem actually gives the following a priori stronger result:

**Theorem 3.1. (Hammersley-Clifford Theorem, strong version)** *Let the  $X \subset \mathcal{A}^{\mathbb{Z}^d}$  be a topological Markov field with a safe symbol. Then:*

- (1) *Any Markov cocycle on  $X$  is a Gibbs cocycle given by a nearest neighbor interaction. In our notation this is expressed by:  $\mathcal{M}_X = \mathcal{G}_X$*
- (2) *Any shift-invariant Markov cocycle on  $X$  is a Gibbs cocycle given by a shift-invariant nearest neighbor interaction. In our notation this is expressed by:  $\mathcal{M}_X^\sigma = \mathcal{G}_X^\sigma$ .*

It is easily verified that any topological Markov field  $X$  which satisfies one of the conclusions of the “strong version” immediately satisfies the corresponding conclusion of the “weak version”. We will demonstrate in the following section that the converse implication is false in general.

For  $X$  with a safe symbol, the “strong” and “weak” versions are easily equivalent because of the following simple claim:



**Proposition 3.2.** *Let the  $X \subset \mathcal{A}^{\mathbb{Z}^d}$  be a topological Markov field with a safe symbol. Then any Markov random field  $\mu$  adapted to  $X$  has  $\text{supp}(\mu) = X$ .*

*Proof.* Let  $\mu$  be a Markov random field adapted to  $X$ . We need to show that for any finite  $F \subset \mathbb{Z}^d$  and any  $a \in \mathcal{B}_F(X)$ ,  $\mu([a]_F) > 0$ . Denote  $\tilde{F} = F \cup \partial F$ . Let  $b \in \mathcal{B}_{\tilde{F}}(X)$  and  $c \in \mathcal{B}_{\partial \tilde{F}}(X)$  satisfy  $\mu([b \vee c]_{\tilde{F} \cup \partial \tilde{F}}) > 0$ . In particular,  $\mu([c]_{\partial \tilde{F}}) > 0$ . Let  $\tilde{b} \in \mathcal{B}_{\tilde{F}}(X)$  be given by

$$\tilde{b}_n = \begin{cases} b_n & n \in F \\ \star & n \in \partial F. \end{cases}$$

The configuration  $\tilde{b} \vee c$  is admissible in  $X$ , because  $\star$  is a safe-symbol. Again, by the safe symbol property it follows that  $\tilde{a} \in \mathcal{B}_{\tilde{F}}(X)$  where:

$$\tilde{a}_n = \begin{cases} a_n & n \in F \\ \star & n \in \partial F. \end{cases}$$

Since  $X$  is a topological Markov field,  $\tilde{a} \vee c \in \mathcal{B}_{\tilde{F} \cup \partial \tilde{F}}(X)$ . Since  $\mu$  is an adapted Markov random field it follows that  $\mu([\tilde{a}]_{\tilde{F}} \mid [c]_{\partial \tilde{F}}) > 0$ , and since  $\mu([c]_{\partial \tilde{F}}) > 0$  it follows that  $\mu([\tilde{a} \vee c]_{\tilde{F} \cup \partial \tilde{F}}) > 0$ , so in particular  $\mu([a]_F) > 0$ .  $\square$

**Remark:** Proposition 3.2 is a particular instance of the more general fact that any  $\Delta_X$ -nonsingular measure  $\mu$  satisfies  $\text{supp}(\mu) = X$ , whenever  $\Delta_X$  is a topologically minimal. The latter condition means the  $\Delta_X$  equivalence class of any  $x \in X$  is dense in  $X$ .

**Remark:** When the underlying graph is  $\mathbb{Z}^d$ , any shift-invariant topological Markov field  $X$  with a safe symbol is actually a nearest-neighbor shift of finite type. This follows using arguments similar to those appearing in the proof of proposition 3.2.

**3.2. The Pivot Property.** We shall now consider a weaker property than that of having a safe symbol. Let  $X$  be a topological Markov field. If  $x, y \in X$  only differ at a single  $n \in \mathbb{Z}^d$ , then the pair  $(x, y)$  will be called a *pivot move* in  $X$ . A subshift  $X$  is said to have the *pivot property* if for all  $(x, y) \in \Delta_X$  such that  $x \neq y$  there exists a finite sequence of points  $x^{(1)} = x, x^{(2)}, \dots, x^{(k)} = y \in X$  such that each  $(x^{(i)}, x^{(i+1)})$  is a pivot move. In this case we say  $x^{(1)} = x, x^{(2)}, \dots, x^{(k)} = y$  is a chain of pivots from  $x$  to  $y$ . Sometimes in the literature the pivot property is called “local move connectedness”. This condition is the same as (3) in Theorem 4.1 of [3]. Here are some examples of subshifts which have the pivot property:

- (1) Any subshift with a trivial homoclinic relation.
- (2) Any subshift with a safe symbol.
- (3) The “3-colored chessboard” (see below).
- (4)  $r$ -colorings of  $\mathbb{Z}^d$  with  $r \geq 2d + 2$ .

**Proposition 3.3.** *Let  $X$  be any subshift with the pivot property. Then the dimension of  $\mathcal{M}_X^\sigma$  is finite.*

*Proof.* Let  $(x, y) \in \Delta_X$ . Let  $x = x^{(1)}, x^{(2)}, x^{(3)}, \dots, x^{(k)} = y$  be a chain of pivots from  $x$  to  $y$ . Then

$$(3.2) \quad M(x, y) = \sum_{i=1}^{k-1} M(x^{(i)}, x^{(i+1)}).$$

If  $x^{(i)}, x^{(i+1)}$  differ only at  $m_i \in \mathbb{Z}^d$ . Then  $M(x^{(i)}, x^{(i+1)}) = M(\sigma_{-m_i}x^{(i)}, \sigma_{-m_i}x^{(i+1)})$  depends only on  $\sigma_{-m_i}x^{(i)}|_{\{0\} \cup \partial\{0\}}$  and  $\sigma_{-m_i}x^{(i+1)}|_{\{0\} \cup \partial\{0\}}$ . Therefore the dimension of the space of shift-invariant Markov cocycles is bounded by  $|\mathcal{B}_{\{(0,0)\} \cup \partial\{(0,0)\}}|^2$ .  $\square$

**Remark:** More generally, the same proof shows that  $\dim(\mathcal{M}_X) < \infty$  whenever there are a finite number of maps  $f_1, \dots, f_k : X \rightarrow X$  and a finite  $F \subset \mathbb{Z}^d$  so that  $(f_i(x))_n = x_n$  whenever  $n \notin F$  and  $\Delta_x$  is the orbit relation generated by

$$\{\sigma_n^{-1} \circ f_i \circ \sigma_n : n \in \mathbb{Z}^d, i = 1, \dots, k\}.$$

This corresponds to subshifts with the “generalized pivot property” or “local move connectedness”: A generalized pivot move from  $x \in X$  to  $y \in X$  is defined when  $x$  and  $y$  differ only in a fixed finite set  $F \subset \mathbb{Z}^d$ . “Domino tiles” are an interesting and well known and interesting example for a subshift with the generalized pivot property [10].

#### 4. $\mathbb{Z}_r$ -HOMOMORPHISMS, 3-COLORED CHESSBOARDS AND HEIGHT FUNCTIONS

Recall that a *graph-homomorphism* from the graph  $G_1 = (V_1, E_1)$  to the graph  $G_2 = (V_2, E_2)$  is a function  $f : V_1 \rightarrow V_2$  from the vertex set of  $G_1$  to the vertex set of  $G_2$  such that  $f$  sends edges in  $G_1$  to edges in  $G_2$ . Namely, if  $(v, w) \in E_1$  then  $(f(v), f(w)) \in E_2$ .

For the purposes of this paper, a *height function* on  $\mathbb{Z}^d$  is graph homomorphism from the standard Cayley graph of  $\mathbb{Z}^d$  to the standard Cayley graph of  $\mathbb{Z}$ . Equivalently, a height function is a function  $\hat{x} \in \mathbb{Z}^{\mathbb{Z}^d}$  such that for any two adjacent sites  $m, n \in \mathbb{Z}^2$ ,  $|\hat{x}_m - \hat{x}_n| = 1$ . Let  $Ht^{(d)} \subset \mathbb{Z}^{\mathbb{Z}^d}$  denote the set of height functions on  $\mathbb{Z}^d$ .

We now introduce a certain family of  $\mathbb{Z}^d$ -shifts of finite type  $X_r^{(d)}$ , where  $r, d \in \mathbb{N}$ , and  $r > 1$ : Denote by  $\mathbb{Z}_r = \mathbb{Z}/r\mathbb{Z} \cong \{0, \dots, r-1\}$  the finite cyclic group of residues modulo  $r$ . Let  $\phi_r : Ht^{(d)} \rightarrow (\mathbb{Z}_r)^{\mathbb{Z}^d}$  be defined by

$$\phi_r(\hat{x})_n = \hat{x}_n \mod r \text{ for } n \in \mathbb{Z}^d.$$

The  $\mathbb{Z}^d$  subshift  $X_r^{(d)}$  is defined by:

$$X_r^{(d)} = \phi_r(Ht^{(d)}).$$

For  $r = 2$  it is easily verified that  $X_2^{(d)}$  consists precisely a two points  $x^{even}, x^{odd}$ . These are “chessboard configurations”, given by  $x_n^{even} = \mathbf{1}_{[\|n\|=0 \mod 2]}$  and  $x_n^{odd} = \mathbf{1}_{[\|n\|=1 \mod 2]}$  where  $\mathbf{1}$  is the indicator function.

For  $r \neq 1, 4$ , there turns out to be a direct and simple interpretation for the subshift  $X_r^{(d)}$  as the set of graph homomorphisms from the standard Cayley graph of  $\mathbb{Z}^d$  to the standard Cayley graph of  $\mathbb{Z}_r$ . This is expressed by the following proposition.

**Proposition 4.1.** *For any  $d \geq 2$  and  $r \in \mathbb{N} \setminus \{1, 4\}$ ,  $X_r^{(d)}$  is a nearest neighbor shift of finite type given by*

$$X_r^{(d)} = \{x \in (\mathbb{Z}_r)^{\mathbb{Z}^d} : x_n - x_m = \pm 1 \mod r, \text{ whenever } \|n - m\|_1 = 1\}.$$

*Proof.* When  $r = 2$ , by our previous remark  $X_2^{(d)} = \{x^{odd}, x^{even}\}$ , and the claim of the proposition is easily verified. Thus, from now assume  $r \in \mathbb{N} \setminus \{1, 2, 4\}$ . Temporarily, let us denote

$$Y_r^{(d)} = \{x \in (\mathbb{Z}_r)^{\mathbb{Z}^d} : x_n - x_m = \pm 1 \mod r, \text{ whenever } \|n - m\|_1 = 1\}.$$

The claim we need to establish is that  $Y_r^{(d)} = X_r^{(d)}$ , where  $X_r^{(d)} = \phi_r(Ht^{(d)})$ . For any  $\hat{x} \in Ht^{(d)}$  and  $m, n \in \mathbb{Z}^d$  with  $\|n - m\|_1 = 1$ , by definition of  $Ht^{(d)}$ , we have  $|\hat{x}_n - \hat{x}_m| = 1$ . Thus,  $\phi_r(\hat{x})_n - \phi_r(\hat{x})_m = \pm 1 \pmod r$ , and so  $\phi_r(\hat{x}) \in Y_r^{(d)}$ . The inclusion  $X_r^{(d)} \subset Y_r^{(d)}$  is established.

It remains to be shown that the map  $\phi_r : Ht^{(d)} \rightarrow Y_r^{(d)}$  is surjective:

Fix some enumeration  $\mathbb{Z}^d = \{n_1, n_2, \dots\}$ , so that  $n_i$  is adjacent to some  $n_j$  for  $j < i$ . A “breadth-first search” of the Cayley graph of  $\mathbb{Z}^d$  is one way to obtain such an enumeration. Given  $x \in X_r^{(d)}$ , set  $\hat{x}_{n_1} = x_{n_1}$ , and recursively define  $\hat{x}_{n_{i+1}} = \hat{x}_{n_i} + [x_{n_{i+1}} - x_{n_i}]$ , where

$$[x_{n_{i+1}} - x_{n_i}] = \begin{cases} 1 & \text{if } x_{n_{i+1}} - x_{n_i} = 1 \pmod r \\ -1 & \text{if } x_{n_{i+1}} - x_{n_i} = -1 \pmod r \end{cases}.$$

To see that  $|\hat{x}_n - \hat{x}_m| = 1$  whenever  $\|n - m\|_1 = 1$ , it is sufficient to prove that the resulting  $\hat{x}$  is independent of the choice of enumeration  $n_1, n_2, \dots$  provided it starts with  $n_1$ . Equivalently, we need to show that whenever  $m_0, m_1, \dots, m_T \in \mathbb{Z}^d$  satisfy  $\|m_k - m_{k-1}\|_1 = 1$  for  $k = 1, \dots, T$  and  $m_0 = m_T$ , we have  $\sum_{k=1}^T [x_{m_k} - x_{m_{k-1}}] = 0$ . We can further reduce this to proving the above for  $T = 4$ , since cycles of length 4 are a basis for all cycles in the Cayley graph of  $\mathbb{Z}^d$  (equivalently, for any two paths  $p$  and  $q$  in the Cayley graph of  $\mathbb{Z}^d$  with common start and end points there is a “discrete homotopy” from one to the other). Thus, the proof reduces to the claim that for all  $x \in X_r^{(d)}$  and any  $i, j \in \{1, \dots, d\}$ :

$$[x_{e_i} - x_0] + [x_{e_i+e_j} - x_{e_i}] = [x_{e_j} - x_0] + [x_{e_i+e_j} - x_{e_j}],$$

where addition and the equality is in  $\mathbb{Z}$ . This equality clearly holds modulo  $r$ . Also, the 2 summands in each side are both in  $\{\pm 1\}$ . Since  $r \in \mathbb{N} \setminus \{1, 2, 4\}$ , whenever  $A, B, C, D \in \{\pm 1\}$  and  $A + B = C + D \pmod r$ , it follows that  $A + B = C + D$  as integers. Thus, the claim holds.  $\square$

In the particular case  $r = 3$ ,  $X_3^{(d)}$  is a presentation of a shift of finite type known as the  $d$ -dimensional 3-colored chessboard. The subshift  $X_3^{(d)}$  is the set of  $\{0, 1, 2\}$ -colorings of  $\mathbb{Z}^d$  in which neighboring sites have distinct color. For simplicity, in the upcoming sections the reader can restrict to the case  $r = 3$  and  $d = 2$  so  $X_3^{(d)}$  is the 2-dimensional 3-colored chessboard.

The essential part of this proof, in a slight reformulation amounts to proving that a certain shift-cocycle over  $X_r^{(d)}$  is well defined, as in [20, 22]. Similar methods will be employed to realize the basis of the space of shift-invariant Markov cocycles on the spaces  $X_r^{(d)}$  in section 5.

We remark that in the exceptional case  $r = 4$ , the same arguments show that  $X_4^{(d)}$  is still a shift of finite type, although not nearest neighbor, and we have:

$$\begin{aligned} X_4^{(d)} &= \{x \in (\mathbb{Z}_r)^{\mathbb{Z}^d} \mid x_n - x_m = \pm 1 \pmod 4, \text{ whenever } \|n - m\|_1 = 1 \\ &\quad \text{and } [x_0 - x_{e_i}] + [x_{e_i} - x_{e_i+e_j}] = [x_0 - x_{e_j}] + [x_{e_j} - x_{e_i+e_j}] \\ &\quad \text{for all } i \text{ and } j\}. \end{aligned}$$

**Lemma 4.2.** Fix any  $d \geq 2$  and  $r \in \mathbb{N} \setminus \{1, 2\}$ . For  $x \in X_r^{(d)}$  any two pre-images under  $\phi_r$  differ by a constant integer multiple of  $r$ , that is, if  $\phi_r(\hat{x}^{(1)}) = \phi_r(\hat{x}^{(2)})$  then there exists  $M \in \mathbb{Z}$  so that  $\hat{x}_n^{(1)} - \hat{x}_n^{(2)} = rM$  for all  $n \in \mathbb{Z}^d$ .

*Proof.* As in the proof of proposition 4.1, fix some enumeration  $\mathbb{Z}^d = \{n_1, n_2, \dots\}$ , so that  $n_i$  is adjacent to some  $n_j$  for  $j < i$ .

The proof that any two pre-images differ by a constant integer multiple of  $r$  can be carried out by a simple induction: Induct on  $k \in \mathbb{N}$  to show that if  $\phi_r(\hat{x}) = \phi_r(\hat{y})$  then

$$\hat{x}_{n_k} - \hat{y}_{n_k} = \hat{x}_{n_1} - \hat{y}_{n_1} \in r\mathbb{Z}.$$

□

In the following, to avoid cumbersome superscripts, we will fix some dimension  $d \geq 2$ , and denote  $Ht := Ht^{(d)}$ ,  $X_r^{(d)} := X_r$ .

By following the arguments in the proof of lemma 4.2 it follows that if  $\hat{x}, \hat{y} \in Ht$  such that  $\hat{x}_{n_0} = \hat{y}_{n_0}$  for some  $n_0 \in \mathbb{Z}^d$ . Then  $\hat{x} = \hat{y}$  on the connected component of  $\{m | \phi_r(\hat{x})_m = \phi_r(\hat{y})_m\}$  containing  $n_0$ . Note that if two height functions  $\hat{x}, \hat{y} \in Ht$  differ at a unique  $n \in \mathbb{Z}^d$  then  $|\hat{x}_n - \hat{y}_n| = 2$ .

**Corollary 4.3.** For any  $(x, y) \in \Delta_{X_r}$ , there exist  $(\hat{x}, \hat{y}) \in \Delta_{Ht}$  so that  $\phi_r(\hat{x}) = x$  and  $\phi_r(\hat{y}) = y$ . The differences  $\hat{x}_n - \hat{y}_n$  are an integer multiple of  $r$  independent of the chosen  $(\hat{x}, \hat{y}) \in \Delta_{Ht}$  for all  $n \in \mathbb{Z}^d$  where  $x_n = y_n$ . Also, for all  $n \in \mathbb{N}$ ,  $(\hat{x}_n - \hat{y}_n) \in 2\mathbb{Z}$ . In particular if for  $x, y \in X_r$ ,  $x = y$  at all sites except some given site  $n_0 \in \mathbb{Z}^d$ , then there exists  $\hat{x}, \hat{y} \in Ht$  such that  $\phi_r(\hat{x}) = x$ ,  $\phi_r(\hat{y}) = y$  and

$$|\hat{x}_n - \hat{y}_n| = \begin{cases} 2 & \text{if } n = n_0 \\ 0 & \text{otherwise} \end{cases}.$$

It is a known and useful fact that the 3-colored chessboard has the pivot property. This can be shown, for instance, using height functions. Essentially the same argument shows that  $X_r$  has the pivot property. We include a proof below, both for completeness and because similar arguments appear in the proofs of certain claims in the subsequent sections:

**Proposition 4.4.** For any  $d \geq 2$  and  $r \in \mathbb{N}$  the subshift  $X_r$  has the pivot property. In other words, given any  $(x, y) \in \Delta_{X_r}$  there exist points  $x = z^{(0)}, z^{(1)}, \dots, z^{(N)} = y \in X_r$  such that for all  $0 \leq k < N$ , there is a unique  $n_k \in \mathbb{Z}^d$  for which  $z_{n_k}^{(k)} \neq z_{n_k}^{(k+1)}$ .

*Proof.* By corollary 4.3, given  $(x, y) \in \Delta_{X_r}$  choose homoclinic height functions  $(\hat{x}, \hat{y}) \in \Delta_{Ht}$  with  $\phi_r(\hat{x}) = x$  and  $\phi_r(\hat{y}) = y$ . We will proceed by induction on  $D = \sum_{n \in \mathbb{Z}^d} |\hat{x}_n - \hat{y}_n|$ . Note that by corollary 4.3 this is well defined, that is, the differences  $(\hat{x}_n - \hat{y}_n)$  in the sum do not depend on the choice of  $(\hat{x}, \hat{y})$ . When  $D = 0$ , then  $x = y$  and the claim is trivial. Now, suppose  $D > 0$ . Let

$$\begin{aligned} F_+ &= \{n \in \mathbb{Z}^d : (\hat{x}_n - \hat{y}_n) > 0\} \\ F_- &= \{n \in \mathbb{Z}^d : (\hat{x}_n - \hat{y}_n) < 0\}, \end{aligned}$$

that is,  $F_+ \subset \mathbb{Z}^d$  is the finite set of sites where  $\hat{x}$  is strictly above  $\hat{y}$  and  $F_- \subset \mathbb{Z}^d$  is the finite set of sites where  $\hat{y}$  is strictly above  $\hat{x}$ . Without loss of generality assume that  $F_+$  is non-empty. Since  $(\hat{x}_n - \hat{y}_n) \in 2\mathbb{Z}$ , for all  $n \in F_+$ ,  $\hat{x}_n - \hat{y}_n \geq 2$ . Let  $n_0 \in F_+$  be the site such that  $\hat{x}_{n_0} = \max\{\hat{x}_n : n \in F_+\}$ . It follows that  $\hat{x}_{n_0} - \hat{x}_m = 1$  for all  $m$  neighboring  $n_0$ . We

can thus define  $\hat{z} \in Ht$  which is equal to  $\hat{x}$  everywhere except at  $n_0$ , where  $\hat{z}_{n_0} = \hat{x}_{n_0} - 2$ . Now set  $z^{(1)} = \phi_r(\hat{z})$  and apply the induction hypothesis on  $(z^{(1)}, y)$ .  $\square$

## 5. MARKOV COCYCLES ON $X_r$

Our goal in the current section is to describe the space of shift-invariant Markov cocycles on  $X_r$ , when  $r \in \mathbb{N} \setminus \{1, 2\}$ , and the sub-space of Gibbs-cocycles for shift-invariant nearest neighbor interactions.

**Lemma 5.1.** *Let  $F$  be a finite set  $F \subset \mathbb{Z}^d$  and  $x, y, z, w \in X_r$  such that for all  $n \in F^c$ ,  $x_n = y_n$  and  $z_n = w_n$  and for all  $n \in F \cup \partial F$ ,  $x_n = z_n$  and  $y_n = w_n$ . Let  $(\hat{x}, \hat{y}) \in \Delta_{Ht}$  and  $(\hat{z}, \hat{w}) \in \Delta_{Ht}$  be homoclinic pairs of height functions projecting via  $\phi_r$  to  $(x, y) \in \Delta_{X_r}$  and  $(z, w) \in \Delta_{X_r}$  respectively. Then  $\hat{x}_n - \hat{y}_n = \hat{z}_n - \hat{w}_n$  for all  $n \in \mathbb{Z}^d$ .*

*Proof.* Let  $F_0 \subset \mathbb{Z}^d$  denote the infinite connected component of  $F^c$ . For  $n \in F_0$ , we clearly have  $\hat{x}_n - \hat{y}_n = \hat{z}_n - \hat{w}_n = 0$ . We can now prove the required by induction on the distance from  $n \in \mathbb{Z}^d$  to  $F_0$  that  $\hat{x}_n - \hat{y}_n = \hat{z}_n - \hat{w}_n$ . Given  $n \in \mathbb{Z}^d \setminus F_0$ , find a neighbor  $m$  of  $n$  which is closer to  $F_0$ . By the induction hypothesis,  $\hat{x}_m - \hat{y}_m = \hat{z}_m - \hat{w}_m$ .

If either  $n \in F$  or  $m \in F$ , then both  $m$  and  $n$  are in  $F \cup \partial F$  and so  $\hat{x}_n - \hat{x}_m = \hat{z}_n - \hat{z}_m$  and  $\hat{y}_n - \hat{y}_m = \hat{w}_n - \hat{w}_m$ , which implies  $\hat{x}_n - \hat{x}_m = \hat{z}_n - \hat{z}_m$  and  $\hat{y}_n - \hat{y}_m = \hat{w}_n - \hat{w}_m$ . Subtracting the equations and applying the induction hypothesis, we conclude in this case that  $\hat{x}_n - \hat{y}_n = \hat{z}_n - \hat{w}_n$ .

Otherwise,  $n, m \in F^c$  and so  $\hat{x}_n - \hat{x}_m = \hat{y}_n - \hat{y}_m$  and  $\hat{z}_n - \hat{z}_m = \hat{w}_n - \hat{w}_m$ , and again we conclude that  $\hat{x}_n - \hat{y}_n = \hat{z}_n - \hat{w}_n$ .  $\square$

For  $i \in \mathbb{Z}_r$  and integers  $a, b$  with  $a - b \in 2\mathbb{Z}$ , let

$$N_i(a, b) = \begin{cases} |\{m \in (2\mathbb{Z} + a) \cap (r\mathbb{Z} + i) : m \in [a, b]\}| & \text{if } a \leq b \\ -N_i(b, a) & \text{otherwise} \end{cases}$$

Here  $N_i$  is the “net” number of crossings from  $(i + r\mathbb{Z})$  to  $(i + 2 + r\mathbb{Z})$  in a path going from  $a$  to  $b$  in steps of magnitude 2. Note that

$$(5.1) \quad N_i(a, b) = N_i(a + rn, b + rn)$$

for all  $a, b, n \in \mathbb{Z}$  and

$$(5.2) \quad N_i(a, b) = N_i(a + c, b + c)$$

for all  $i \in \{0, 1, \dots, r-1\}$  and  $a, b, c \in \mathbb{Z}$  such that  $a - b \in r\mathbb{Z} \cap 2\mathbb{Z}$ .

**Proposition 5.2.** *For any  $r \in \mathbb{N} \setminus \{1, 2\}$ , the space  $\mathcal{M}_{X_r}^\sigma$  of shift-invariant Markov cocycles on  $X_r$  has the linear basis:*

$$\{M_0, M_1, \dots, M_{r-1}\},$$

where the cocycle  $M_i$  is given by

$$(5.3) \quad M_i(x, y) = \sum_{n \in \mathbb{Z}^d} N_i(\hat{x}_n, \hat{y}_n),$$

with  $(\hat{x}, \hat{y}) \in \Delta_{Ht}$  height functions projecting to  $(x, y)$  via  $\phi_r$ . In particular,  $\dim \mathcal{M}_{X_r}^\sigma = r$ .

*Proof.* By Corollary 4.3 the function  $M_i$ 's are independent of the choice of corresponding height functions and hence are well defined. We will first show that for  $i = 0, \dots, r-1$ ,  $M_i$  is indeed a shift-invariant Markov cocycle. Since  $N_i(a, c) = N_i(a, b) + N_i(b, c)$  whenever  $a \equiv b \equiv c \pmod{2}$ , it follows that  $M_i(x, z) = M_i(x, y) + M_i(y, z)$  whenever  $x, y, z \in X_r$  are homoclinic so  $M_i$  is a  $\Delta_{X_r}$ -cocycle. Clearly,  $M_i$  is shift-invariant.

It remains to show that  $M_i$  satisfies the Markov property. This is equivalent to showing that  $M_i(x, y) = M_i(z, w)$  whenever  $x, y, z, w \in X_r$  satisfy the assumption in lemma 5.1. In this case by lemma 5.1,  $\hat{x}_n - \hat{y}_n = \hat{z}_n - \hat{w}_n$  for all  $n \in \mathbb{Z}^d$ . Also note that for any  $n \in \mathbb{Z}^d$ , either  $x_n = z_n$  and  $y_n = w_n$  in which case  $\hat{x}_n - \hat{z}_n = \hat{y}_n - \hat{w}_n \in r\mathbb{Z}$  or  $x_n = y_n$  and  $z_n = w_n$ , in which case  $\hat{x}_n - \hat{y}_n = \hat{z}_n - \hat{w}_n \in r\mathbb{Z} \cap 2\mathbb{Z}$ . By equations 5.1 and 5.2 in either case  $N_i(\hat{x}_n, \hat{y}_n) = N_i(\hat{z}_n, \hat{w}_n)$  and summing over all  $n$ 's, indeed  $M_i(x, y) = M_i(z, w)$  as required.

To complete the proof we need to show that any shift-invariant Markov cocycle on  $X_r$  can be uniquely written as a linear combination of  $M_0, \dots, M_{r-1}$ . For  $i \in \{0, \dots, r-1\}$  let  $(x^{(i)}, y^{(i)}) \in \Delta_{X_r}$  be homoclinic points such that  $x_0^{(i)} = i$ ,  $y_0^{(i)} = i+2 \pmod{r}$  and  $x_n^{(i)} = y_n^{(i)}$  for all  $n \in \mathbb{Z}^d \setminus \{0\}$ . Given a shift-invariant Markov cocycle  $M$ , let  $\alpha_i := M(x^{(i)}, y^{(i)})$ . We claim that for this choice,  $M = \sum_{i=0}^{r-1} \alpha_i \cdot M_i$ . Since  $X_r$  has the pivot property (proposition 4.4), by equation 3.2 it is sufficient to show that for any  $(x, y) \in \Delta_{X_r}$  which differ only at a single site ,

$$M(x, y) = \sum_{i=0}^{r-1} \alpha_i \cdot M_i(x, y).$$

By shift-invariance of  $M$  and the  $M_i$ 's it is further enough to show this for  $(x, y)$  which differ only at the origin. In this case, it follows since  $(x, y)$  coincide with  $(x^{(i)}, y^{(i)})$  on the sites  $\{0\} \cup \partial\{0\}$  for some  $i$ , and since  $M$  and the  $M_i$ 's are Markov cocycles we have  $M(x, y) = M(x^{(i)}, y^{(i)}) = \alpha_i$  and

$$\sum_{j=0}^{r-1} \alpha_j \cdot M_j(x, y) = \sum_{j=0}^{r-1} \alpha_j M_j(x^{(i)}, y^{(i)}) = \sum_{j=0}^{r-1} \alpha_j \delta_{i,j} = \alpha_i.$$

□

**Remark:** Without the assumption of shift-invariance, a similar argument shows that any Markov cocycles on  $X_r$  is of the following form:

$$M(x, y) = \sum_{i=0}^{r-1} \sum_{n \in \mathbb{Z}^d} a_{i,n} N_i(\hat{x}, \hat{y}) \text{ with } a_{i,n} \in \mathbb{R} \text{ for all } n \in \mathbb{Z}^d, 1 \leq i \leq r.$$

We now describe the space  $\mathcal{G}_{X_r}^\sigma$  of Gibbs-cocycles corresponding to shift-invariant nearest neighbor interactions for  $X_r$ .

**Proposition 5.3.** *A shift-invariant Markov cocycle on  $X_r$  is a Gibbs cocycle corresponding to a shift-invariant nearest neighbor interaction if and only if it is of the form  $M = \sum_{i=0}^{r-1} \alpha_i M_i$ , with  $\sum_{i=0}^{r-1} \alpha_i = 0$  and  $M_i$ 's as in proposition 5.2. In other words,*

$$\mathcal{G}_{X_r}^\sigma = \left\{ \sum_{i=0}^{r-1} \alpha_i M_i : \sum_{i=0}^{r-1} \alpha_i = 0 \right\}.$$

In particular,  $\dim(\mathcal{G}_{X_r}^\sigma) = r - 1$ .

*Proof.* Let  $M$  be a Gibbs cocycle given by a shift-invariant nearest neighbor interaction  $\phi$ . Choose  $(x^{(i)}, y^{(i)}) \in \Delta_{X_r}$  as in the proof of proposition 5.2. Expanding the Gibbs cocycle by definition we have:

$$\begin{aligned} M(x^{(i)}, y^{(i)}) &= \phi([i+2]_0) - \phi([i]_0) \\ &+ \sum_{j=1}^d (\phi([i+2, i+1]_j) - (\phi([i, i+1]_j) + \phi([i+1, i+2]_j) - \phi([i+1, i]_j))). \end{aligned}$$

Now summing these equations over  $i$  we get:

$$\sum_{i=0}^{r-1} \alpha_i = \sum_{i=0}^{r-1} M(x^{(i)}, y^{(i)}) = 0$$

Conversely, for any values  $\alpha_i = M(x^{(i)}, y^{(i)})$  such that  $\sum_{i=0}^{r-1} \alpha_i = 0$  it is easy to see that there is a solution to the linear equations on the values of  $\phi$ . For instance, set  $\phi([i, i+1]_1) = -\sum_{k=i}^{r-1} \alpha_k$ ,  $\phi([i, i+1]_j) = 0$  for  $j = 2, \dots, d$  and  $\phi([i+1, i]_j) = \phi([i]_0) = 0$  for  $j = 1, \dots, d$  and  $i = 0, \dots, r-1$ .  $\square$

Let  $\hat{M} : \Delta_{X_r} \rightarrow \mathbb{R}$  be the Markov cocycle given by

$$(5.4) \quad \hat{M}(x, y) = \sum_{n \in \mathbb{Z}^d} \hat{y}_n - \hat{x}_n,$$

where  $(\hat{x}, \hat{y})$  satisfy  $\phi_r(\hat{x}) = x$  and  $\phi_r(\hat{y}) = y$ . It is easily verified that  $\hat{M} = \sum_{i=0}^{r-1} M_i$ .

**Corollary 5.4.** *Any shift-invariant Markov cocycle  $M$  on  $X_r$  can be uniquely written as*

$$M = M_0 + \alpha \hat{M}$$

where  $M_0$  is some Gibbs cocycle,  $\alpha \in \mathbb{R}$  and  $\hat{M}$  is given by (5.4).

Thus, the conclusion of the second part of the strong version of the Hammersley-Clifford theorem, regarding shift-invariant Markov cocycles fails for  $X_r$ . Our next proposition asserts that the conclusion of the first part of the strong version of the Hammersley-Clifford theorem still holds for  $X_r$ . This immediately implies the conclusion of the first part of the weak version of the Hammersley-Clifford theorem of  $X_r$ .

**Proposition 5.5.**  $(\mathcal{M}_{X_r} = \mathcal{G}_{X_r})$

Let  $M : \Delta_{X_r} \rightarrow \mathbb{R}$  be a Markov cocycle. There exists a nearest neighbor interaction  $\phi$ , which is not necessarily shift-invariant, so that  $M$  is given by the interaction  $\phi$ .

*Proof.* Given  $M \in \mathcal{M}_{X_r}$ , we will define a compatible nearest neighbor interaction  $V$ . For  $n = (n_1, \dots, n_d) \in \mathbb{Z}^d$ , and  $1 \leq j \leq d$ , let  $\phi_{n,j}(a, b)$  denote the weight the interaction  $\phi$  assigns the configuration  $(a, b)$  on the edge  $(n, n + e_j)$ . Set  $\phi_{n,j}(a, b) = 0$  whenever  $j \in \{2, \dots, d\}$ . For  $n = (n_1, \dots, n_d) \in \mathbb{Z}^d$  and  $i \in \mathbb{Z}_r$  define

$$\begin{aligned} \phi_{n,1}(i, i+1) &= \begin{cases} 0 & n_1 \leq 0 \\ M(\sigma_n y^{(i)}, \sigma_n x^{(i)}) + \phi_{n-e_1,1}(i+1, i+2) & n_1 > 0 \end{cases} \\ \phi_{n,1}(i+1, i) &= \begin{cases} 0 & n_1 \geq 0 \\ M(\sigma_{n+e_1} y^{(i)}, \sigma_{n+e_1} x^{(i)}) + \phi_{n+e_1,1}(i+2, i+1) & n_1 < 0 \end{cases} \end{aligned}$$

where as in the proof of proposition 5.2,  $(x^{(i)}, y^{(i)}) \in \Delta_{X_r}$  are such that  $x_0^{(i)} = i \bmod r$ ,  $y_0^{(i)} = i + 2 \bmod r$  and  $x_n^{(i)} = y_n^{(i)}$  for all  $n \in \mathbb{Z}^d \setminus \{0\}$ . This is well defined by induction on  $|n_1|$ . By direct computation it follows that

$$(5.5) \quad M(x, y) = \sum_{n \in \mathbb{Z}^d} \sum_{j=1}^d \phi_{n,j}(x_n, x_{n+e_j}) - \phi_{n,j}(y_n, y_{n+e_j})$$

whenever  $x$  and  $y$  differ only in a single site. Note that the expressions in both sides of (5.5) are  $\Delta_{X_r}$ -cocycles. Thus by the pivot property of  $X_r$  (proposition 4.4) the equality in (5.5) holds for any  $(x, y) \in \Delta_{X_r}$ .  $\square$

Combining the above results we obtain:

**Corollary 5.6.** *There exists a shift-invariant Markov cocycles on  $X_r$  which is given by a nearest interaction but not by a shift-invariant nearest neighbor interaction. That is,*

$$\mathcal{G}_{X_r}^\sigma \neq \mathcal{G}_{X_r} \cap M_{X_r}^\sigma.$$

## 6. MARKOV RANDOM FIELDS ON $X_r$ ARE GIBBS

Our main goal is to prove the following result:

**Theorem 6.1.** *Any shift-invariant Markov random field adapted to  $X_r$  is a Gibbs state for some shift-invariant nearest neighbor interaction. In particular any shift-invariant Markov random field such that its support is  $X_r$  is a Gibbs state for some shift-invariant nearest neighbor interaction.*

Theorem 6.1 implies that the conclusion of the second part of the weak version of Hammersley-Clifford theorem holds for  $X_r$ , although the argument is very different from the safe-symbol case.

For a subshift  $X$ , a point  $x \in X$  will be called *frozen* if its homoclinic class is singleton. This notion coincides with the notion of frozen coloring in [3]. By proposition 4.4,  $X_r$  has the pivot property so  $x \in X_r$  is frozen if and only if for every  $n \in \mathbb{Z}^d$ ,  $x_j \neq x_k$  for some  $j, k \in \partial\{n\}$ , that is, any site is adjacent to at least two sites with distinct symbols. A subshift  $X$  will be called *frozen* if it consists of frozen points, equivalently  $\Delta_X$  is the diagonal. A measure on  $X$  will be called *frozen* if its support consists of frozen points. Note that the collection of frozen points of a given topological Markov field  $X$  are themselves a topological Markov field.

We derive theorem 6.1 as an immediate corollary of the following proposition:

**Proposition 6.2.** *Let  $\mu$  be a shift-invariant Markov random field adapted to  $X_r$  with Random-Nikodym cocycle equal to the restriction  $e^M$  on its support where  $M \in \mathcal{M}_{X_r}^\sigma \setminus \mathcal{G}_{X_r}^\sigma$  is a Markov cocycle which is not given by a shift-invariant nearest neighbor interaction. Then  $\mu$  is frozen.*

Note that any frozen probability measure is Gibbs with a nearest neighbor interaction because the homoclinic relation of the support of the measure is trivial. The intuition behind the proof of this proposition is the following. For a Markov cocycle  $M = \sum_{i=1}^r \alpha_i M_i$  the condition  $\sum_{i=1}^r \alpha_i > 0$  indicates an inclination to raise the height function. However  $\sigma$ -invariance implies the existence of a well defined “slope” for the height function in any direction. Unless this slope is extremal (equivalently, maximal in some direction), this leads to a contradiction.



In preparation for the proof, we set up some auxiliary results.

**6.1. Real valued cocycles for measure-preserving  $\mathbb{Z}^d$ -actions.** We momentarily pause our discussion about Markov random fields on  $X_r$  to discuss cocycles for measure-preserving  $\mathbb{Z}^d$  actions. Let  $(X, \mathcal{F}, \mu, T)$  be an ergodic measure-preserving  $\mathbb{Z}^d$  action. A measurable function  $c : X \times \mathbb{Z}^d \rightarrow \mathbb{R}$  is called a  $T$ -cocycle if it satisfies the following equation  $\mu$ -almost everywhere:

$$(6.1) \quad c(x, n + m) = c(x, n) + c(T_n x, m).$$

The relevance to our context is as follows: taking  $T = \sigma$  to be the shift action, the function  $c : X_r \times \mathbb{Z}^d \rightarrow \mathbb{R}$  given by

$$c(x, n) = \hat{x}_n - \hat{x}_0 \text{ where } \hat{x} \in Ht \text{ and } \phi(\hat{x}) = x,$$

satisfies equation (6.1), that is, the difference of heights between various sites gives us a  $\mathbb{Z}^d$ -cocycle.

We will use the following lemma:

**Lemma 6.3.** *Let  $(X, \mathcal{F}, \mu, T)$  be an ergodic measure-preserving  $\mathbb{Z}^d$  action and  $c : X \times \mathbb{Z}^d \rightarrow \mathbb{R}$  be a measurable cocycle such that for any  $n \in \mathbb{Z}^d$  the function  $f_{c, \vec{n}}(x) := c(x, \vec{n})$  is in  $L^1(\mu)$ , then for any  $n = (n_1, n_2, \dots, n_d) \in \mathbb{Z}^d$*

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{c(x, k \cdot n)}{k} &= \int c(x, n) d\mu(x) \\ &= \sum_{i=1}^d n_i \int c(x, e_i) d\mu(x). \end{aligned}$$

*The convergence holds almost everywhere with respect to  $\mu$  and also in  $L^1(\mu)$ .*

*Proof.* By the cocycle equation (6.1) we have:

$$c(x, k \cdot n) = \sum_{i=0}^{k-1} c(T_n^i x, n).$$

The existence of almost everywhere and  $L^1$  limit  $\bar{f}(x) := \lim_{k \rightarrow \infty} \frac{c(x, k \cdot n)}{k}$  follows from the pointwise and  $L^1$  ergodic theorems. It remains to show that the limit is constant. We do this by showing that the limit is  $T$ -invariant: For any  $m \in \mathbb{Z}^d$  we have:

$$c(x, kn) = c(x, m) + c(T_m x, kn) + c(T_{m+kn} x, -m).$$

Thus,

$$|\bar{f}(x) - \bar{f}(T_m(x))| \leq \lim_{k \rightarrow \infty} \frac{1}{k} (|c(x, m)| + |c(T_{m+kn} x, -m)|) \leq \lim_{k \rightarrow \infty} \frac{2\|m\|_1}{k} = 0.$$

Therefore,

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{c(x, k \cdot n)}{k} &= \int c(x, n) d\mu(x) \\ &= \sum_{i=1}^d n_i \int c(x, e_i) d\mu(x). \end{aligned}$$

□

**Remark:** In the specific case that  $T$  is *totally ergodic*, meaning that the individual action of each  $T_n$  is ergodic for any  $n \in \mathbb{Z}^d \setminus \{0\}$ , the lemma above is completely obvious since  $\frac{c(x,k \cdot n)}{k} = \frac{1}{k} \sum_{j=0}^{k-1} c(T_n^j x, n)$ , which is an ergodic average.

**Corollary 6.4.** *Let  $\mu$  be an ergodic measure on  $X_r$ , and  $n = (n_1, \dots, n_d) \in \mathbb{R}^d$ . Then  $\mu$ -almost surely*

$$\lim_{k \rightarrow \infty} \frac{1}{k} (\hat{x}_{[kn]} - \hat{x}_0) = \sum_{i=1}^d n_i v_i$$

where for  $w = (w_1, \dots, w_d) \in \mathbb{R}^d$ ,  $[w] = ([w_1], \dots, [w_d]) \in \mathbb{Z}^d$ ,  $\hat{x} \in Ht$  satisfies  $\phi_r(\hat{x}) = x$  and

$$v_j = \mu(\{x \in X_r : x_{e_j} - x_0 = 1 \pmod{r}\}) - \mu(\{x \in X_r : x_{e_j} - x_0 = -1 \pmod{r}\})$$

for  $j = 1, \dots, d$ .

**6.2. Maximal height functions.** For  $\hat{x} \in Ht$  and a finite  $F \subset \mathbb{Z}^d$ , let

$$Ht_{\hat{x}, F} = \{\hat{y} \in Ht \mid \hat{y}_n = \hat{x}_n \text{ if } n \notin F\}.$$

Note that if  $\phi_r(\hat{x}) = x$  then  $Ht_{\hat{x}, F}$  is in bijection with  $\{y \in X_r \mid y_n = x_n \text{ if } n \notin F\}$  via the map  $\phi_r$ . Consider the partial ordering on  $Ht_{\hat{x}, F}$  given by  $\hat{y} \geq \hat{z}$  if  $\hat{y}_n \geq \hat{z}_n$  for all  $n \in \mathbb{Z}^d$ .

**Lemma 6.5.** *Let  $\hat{x} \in Ht$  and  $N \in \mathbb{N}$  be given. Then  $(Ht_{\hat{x}, F}, \geq)$  has a maximum. If the maximum is attained by the height function  $\hat{y}$  then for all  $n \in F$ :*

$$\hat{y}_n = \min\{\hat{x}_k + \|n - k\|_1 \mid k \in \partial F\}.$$

*Proof.* Given height functions  $\hat{y}, \hat{z} \in Ht(\hat{x}, F)$ , note that the function  $\hat{w}$  defined by

$$\hat{w}_i = \max(\hat{y}_i, \hat{z}_i).$$

is an element of  $Ht_{\hat{x}, F}$ .

To see why  $\hat{w}$  is a valid height function consider adjacent sites  $i$  and  $j \in \mathbb{Z}^d$ . Then either

$$\hat{y}_i \geq \hat{z}_i \quad , \quad \hat{y}_j \geq \hat{z}_j$$

or

$$\hat{y}_i \leq \hat{z}_i \quad , \quad \hat{y}_j \leq \hat{z}_j.$$

In either case  $|\hat{w}_i - \hat{w}_j| = 1$ . This proves that  $\hat{w}$  is a valid height function. Also  $\hat{w}_i = \max(\hat{y}_i, \hat{z}_i) = \hat{x}_i$  for all  $i \in F^c$ . Hence  $\hat{w} \in Ht_{\hat{x}, F}$ . Since  $Ht_{\hat{x}, F}$  is finite, it has a maximum.

Suppose the maximum is attained by a height function  $\hat{y}$ . Let  $i \in F$ ,  $k \in \partial F$  and  $(i_1 = i), i_2, i_3, \dots, i_p, (i_{p+1} = k)$  be a shortest path between  $i$  and  $k$ . Then

$$\hat{y}_i = \sum_{t=1}^p \hat{y}_{i_t} - \hat{y}_{i_{t+1}} + \hat{y}_k = \sum_{t=1}^p \hat{y}_{i_t} - \hat{y}_{i_{t+1}} + \hat{x}_k.$$

Therefore  $\hat{y}_i \leq \|i - k\|_1 + \hat{x}_k$  which proves that

$$(6.2) \quad \hat{y}_i \leq \min\{\hat{x}_k + \|i - k\|_1 \mid k \in \partial F\}.$$

For proving the reverse inequality, note that if  $\hat{y}$  has a local minimum at some  $n \in F$  then the height at  $n$  can be increased. Since  $\hat{y}$  is the maximum height function, for each  $n \in F$  at least one of the adjacent sites  $m$  satisfy  $\hat{y}_n - \hat{y}_m = 1$ . Iterating this argument, for any  $i \in F$ ,

we can choose a path  $j_1, j_2, j_3, \dots, j_{p+1}$ , with  $j_1 = i$ ,  $j_2, \dots, j_p \in F$ ,  $j_{p+1} \in \partial F$  along which  $\hat{y}$  is increasing:  $\hat{y}_{j_t} - \hat{y}_{j_{t+1}} = 1$  for all  $t \in \{1, 2, \dots, p\}$ . Then

$$\hat{y}_i = \sum_{t=1}^p \hat{y}_{j_t} - \hat{y}_{j_{t+1}} + \hat{y}_{j_{p+1}} \geq \|i - j_{p+1}\|_1 + \hat{y}_{j_{p+1}}.$$

Combining with the inequality 6.2, we get

$$\hat{y}_i = \min\{\hat{x}_k + \|i - k\|_1 \mid k \in \partial F\}.$$

□

Consider a shift-invariant Markov cocycle  $M \in \mathcal{M}_{X_r}^\sigma$ . Recall that by corollary 5.4 there exists  $\alpha \in \mathbb{R}$  and a Gibbs cocycle  $M_0 \in \mathcal{G}_{X_R}^\sigma$  compatible with a shift-invariant nearest neighbor interaction such a that  $M = M_0 + \alpha \hat{M}$ . The following lemma is based on the idea that in the “non-Gibbsian” case  $\alpha \neq 0$ , whenever  $\hat{y}$  is much bigger than  $\hat{x}$ ,  $M(x, y)$  is roughly  $\alpha$  times the volume of  $d + 1$ -dimensional shape bounded between  $\hat{y}$  and  $\hat{x}$ .

As before, for  $N \in \mathbb{N}$ , let  $D_N = \{n \in \mathbb{Z}^d \mid \|n\|_1 \leq N\}$  be the  $N$ -radius  $l^1$ -ball in  $\mathbb{Z}^d$  centered at the origin.

**Lemma 6.6.** *Let  $M = M_0 + \alpha \hat{M}$  be a shift-invariant Markov cocycle on  $X_r^{(d)}$  where  $M_0 \in \mathcal{G}_{X_R}^\sigma$  is a Gibbs cocycle compatible with a shift-invariant nearest neighbor interaction,  $\hat{M}$  is the Markov cocycle given by formula (5.4) and  $\alpha > 0$ . Then there exist a positive constant  $c_1 > 0$  (depending only on  $d$ ) and another positive constant  $c_2 > 0$  (depending only on  $d$  and  $M_0$ ) such that for all  $N \in \mathbb{N}$  and all  $(x, y) \in \Delta_{X_r}$  satisfying  $x_i = y_i$  for  $i \in \mathbb{Z}^d \setminus D_N$  and  $(\hat{x}, \hat{y}) \in \Delta_{Ht}$  for which  $\hat{x} \leq \hat{y}$  we have:*

$$M(x, y) \geq c_1 \alpha (\hat{y}_0 - \hat{x}_0)^{d+1} - c_2 \cdot N^d,$$

where as usual  $(\hat{x}, \hat{y}) \in \Delta_{Ht}$  are height functions corresponding to  $(x, y)$ .

*Proof.* Let  $M = M_0 + \alpha \hat{M}$  be as given in the lemma. We will first show that there exist a suitable constant  $c_1 > 0$  so that  $\hat{M}(x, y) \geq c_1 (\hat{x}_0 - \hat{y}_0)^{d+1}$  for any  $(x, y) \in \Delta_{X_r}$  which satisfy  $\hat{x} \leq \hat{y}$  and  $\hat{y}_0 - \hat{x}_0 > 0$ :

Since  $\hat{y}_i - \hat{x}_i \geq 0$ , we have:

$$\sum_{i \in D_N} (\hat{y}_i - \hat{x}_i) \geq \sum_{j=1}^{\frac{\hat{y}_0 - \hat{x}_0}{2}} \sum_{i \in \partial D_{j-1}} (\hat{y}_i - \hat{x}_i)$$

Since  $\hat{y}_i - \hat{x}_i \geq \hat{y}_0 - \hat{x}_0 - 2\|i\|_1$ :

$$\sum_{j=1}^{\frac{\hat{y}_0 - \hat{x}_0}{2}} \sum_{i \in \partial D_{j-1}} (\hat{y}_i - \hat{x}_i) \geq \sum_{j=1}^{\frac{\hat{y}_0 - \hat{x}_0}{2}} |\partial D_{j-1}| (\hat{y}_0 - \hat{x}_0 - 2(j-1))$$

Setting  $K = \frac{\hat{y}_0 - \hat{x}_0}{2}$ , observe that the sum on the right hand side is roughly the volume of  $d + 1$  dimensional “cone” of height  $K$  and base  $D_K$ . Thus, the sum is proportional (within constants depending only on  $d$ ) to  $K^{d+1}$ . To make this precise, we use the naive estimates

$|\partial D_j| \geq \binom{j+d-1}{d-1} \geq \frac{1}{d!} j^{d-1}$  and

$$\sum_{j=1}^K j^{d-1} (2K - 2j) \geq \int_0^{K-1} x^{d-1} (2K - 2x) dx.$$

This establishes the constant  $c_1 > 0$  so that  $\hat{M}(x, y) \geq c_1 \alpha (\hat{x}_0 - \hat{y}_0)^{d+1}$  (technically, the case  $K = 1$  requires separate consideration, but this is easily handled by reducing the constant  $c_1$ ).

We now show that for any Gibbs cocycle  $M_0$  corresponding to a shift-invariant interaction  $\phi$  there exists a suitable constant  $c_2 > 0$  so that  $|M_0(x, y)| \leq c_2 N^d$  whenever  $(x, y) \in \Delta_{X_r}$  satisfy  $x_n \neq y_n$  for all  $n \in \mathbb{Z}^d \setminus D_N$ : Write

$$M_0(x, y) = \sum_{C \cap D_N \neq \emptyset} \phi(x_C) - \phi(y_C),$$

where  $\phi : \mathcal{B}(X_r) \rightarrow \mathbb{R}$  is a translation invariant nearest neighbor interaction corresponding to  $M_0$ , and the sum is over all cliques  $C$  (edges and vertices) in  $\mathbb{Z}^d$  intersecting  $D_N$ . Therefore

$$|M_0(x, y)| \leq \sum_{n \in D_N} \sum_{C \cap \{n\} \neq \emptyset} |\phi(x_C) - \phi(y_C)|$$

and it follows that  $|M_0(x, y)| \leq c_2 |D_N|$  where  $c'_2 = 2 \sup\{|\sum_{C \cap \{0\} \neq \emptyset} \phi(x_C)| : x \in X_r\}$ . Since  $|D_N| \leq 2^d N^d$ , it follows that  $|M_0(x, y)| \leq c_2 N^d$  with  $c_2 = 2^d c'_2$ .

Putting everything together, we conclude that

$$M(x, y) \geq \alpha \hat{M}(x, y) - |M_0(x, y)| \geq c_1 \alpha (\hat{x}_0 - \hat{y}_0)^{d+1} - c_2 \cdot N^d,$$

as required. □

For the same hypothesis except if  $\alpha < 0$ , we get,

$$M(x, y) \leq c_1 \alpha \cdot K^{d+1} + c_2 N^d$$

for the same constants  $c_1, c_2 > 0$ .

**Lemma 6.7.** *Let  $\mu$  be a shift-invariant measure on  $X_r$  and  $n \in \mathbb{Z}^d$  such that  $\|n\| = 1$ . If*

$$\mu(\{x \in X_r : x_0 - x_n = 1 \pmod{r}\}) = 1,$$

*then  $\mu$  is frozen.*

*Proof.* If  $\mu(\{x \in X_r : x_0 - x_n = 1 \pmod{r}\}) = 1$ , then  $\mu$ -almost surely  $x_{m-n} = x_m + 1 \pmod{r}$  and  $x_{m+n} = x_m - 1$  for all  $m \in \mathbb{Z}^d$ , so  $\mu$ -almost surely  $x$  is frozen. □

In the course of our proof, it will be convenient to restrict to ergodic shift-invariant Markov random fields. The following claim justifies this:

**Theorem 6.8.** *Any shift-invariant Markov random field  $\mu$  with specification  $\Theta$  is in the closure of the convex hull of the ergodic shift-invariant Markov random fields with specification  $\Theta$ .*

*Proof.* See Theorem 14.14 in [11]. □

We now proceed to complete the proof of proposition 6.2.

*Proof.* Since a convex combination of frozen measures is frozen, by theorem 6.8 it suffices to prove that any ergodic Markov random field  $\mu$  adapted to  $X_r$  with its Radon-Nikodym cocycle equal to  $e^M$  on its support where  $M = M_0 + \alpha \hat{M}$  (as in corollary 5.4) and  $\alpha \neq 0$  is frozen.

Choose any  $\mu$  satisfying the above assumptions, assuming without loss of generality that  $\alpha > 0$ . Let

$$v_j = \mu(\{x \in X_r : x_0 - x_{e_j} = 1 \pmod{r}\}) - \mu(\{x \in X_r : x_0 - x_{e_j} = -1 \pmod{r}\})$$

for  $j = 1, \dots, d$ . If  $|v_j| = 1$  for some  $1 \leq j \leq d$ , it follows from lemma 6.7 that  $\mu$  is frozen. We can thus assume that  $|v_j| < 1$  for all  $1 \leq j \leq d$ . Choose  $\epsilon > 0$  satisfying  $\epsilon < \frac{1}{4} \min\{1 - |v_j| : 1 \leq j \leq d\}$ .

Let  $\partial B = \{t \in \mathbb{R}^d : \|t\|_1 = 1\}$ , and choose points  $t^{(1)}, \dots, t^{(M)} \in \partial B$  which are  $\epsilon$ -dense in  $\partial B$  with respect to the norm  $\|\cdot\|_1$ . By this we mean that for any  $t \in \partial B$  there is some  $l \in \{1, \dots, M\}$  so that  $\|t - t^{(l)}\| \leq \epsilon$ . For  $k \in \mathbb{N}$ , let

$$A_k = \bigcap_{l=1}^M \{x \in X_r : |(\hat{x}_0) - \hat{x}_{\lfloor k \cdot t^{(l)} \rfloor} - k \sum_{j=1}^d v_j t_j^{(l)}| < 2\epsilon k\},$$

where for  $v = (v_1, \dots, v_d) \in \mathbb{R}^d$ ,  $\lfloor v \rfloor = (\lfloor v_1 \rfloor, \dots, \lfloor v_d \rfloor) \in \mathbb{Z}^d$ . It follows from corollary 6.4 that for sufficiently large  $k$ ,  $\mu(A_k) > 1 - \epsilon$ . It follows in particular that there exists  $x \in A_k$  such that  $\mu([x]_{\partial D_k}) > 0$  and  $\mu(A_k | [x]_{\partial D_k}) > 1 - 2\epsilon$ . Let  $x$  be as above.

Because  $\{t^{(1)}, \dots, t^{(M)}\}$  are  $\epsilon$ -dense in  $\partial B$ , it follows that

$$(6.3) \quad \hat{y}_0 - \hat{x}_n \leq \sum_{j=1}^d v_j n_j + 3\epsilon k,$$

for any  $y \in A_k \cap [x]_{\partial D_k}$  and  $n \in \partial D_k$ . Choose  $\hat{z}$  which is maximal in  $Ht(\hat{x}, D_k)$ , and let  $z = \phi_r(\hat{z})$ . It follows from lemma 6.5 that for some  $n \in \partial D_k$ ,

$$(6.4) \quad \hat{z}_0 - \hat{x}_n = \|n\|_1.$$

Equations (6.3) and (6.4) together imply that  $\hat{z}_0 - \hat{y}_0 \geq k\epsilon$  for any  $y \in A_k \cap [x]_{\partial D_k}$ . Thus, by lemma 6.6

$$M(y, z) > c_1 \alpha (k \cdot \epsilon)^{d+1} - c_2 k^d > c_3 k^{d+1},$$

the last inequality holding for some  $c_3 > 0$ , when  $k$  is sufficiently large.

It follows that

$$\mu([z]_{F_k} | [x]_{\partial D_k}) \geq \mu([y]_{D_k} | [x]_{\partial D_k}) e^{c_3 k^{d+1}}.$$

Now since  $A_k \cap [x]_{\partial D_k} = \bigcup_y ([y]_{D_k} \cap [x]_{\partial D_k})$ , where the union over all  $y \in \mathcal{B}_{D_k}(X_r)$  such that  $[y]_{D_k} \cap A_k \cap [x]_{\partial D_k} \neq \emptyset$ . Thus  $A_k$  is a union of at most  $|\mathcal{B}_{D_k \cup \partial D_k}(X_r)| = e^{O(k^d)}$  elements. It follows that  $\mu(A_k | [x]_{\partial D_k}) \rightarrow 0$  as  $k \rightarrow \infty$ , in contradiction to our choice of  $x$  which had implied that  $\mu(A_k | [x]_{\partial D_k}) > 1 - 2\epsilon$ .  $\square$

## 7. NON FROZEN ADAPTED SHIFT-INVARIANT MARKOV RANDOM FIELDS ON $X_r$ ARE FULLY-SUPPORTED

We have concluded that any shift-invariant Markov random fields which is adapted with respect to  $X_r$  is a Gibbs measure for some shift-invariant nearest neighbor interaction. Our

next goal is to show that any such measure must be fully-supported on  $X_r$ . For somewhat related results with similar argument, see section 4.3 of [20].

**Proposition 7.1.** *Let  $r \geq 3$  be an odd integer, and let  $\mu$  be a shift-invariant Markov random field adapted with respect to  $X_r$ . Then either  $\text{supp}(\mu) = X_r$  or  $\mu$  is frozen.*

Roughly speaking we shall show that for non-frozen shift-invariant Markov random fields the height function corresponding to a typical point is “not very steep”. Given a height function that is “not very steep”, there is enough flexibility to “deform” the height function while keeping it fixed outside some large ball, in such a way that the restriction of the “deformed” height function in a fixed finite set projects to any word in the language of  $X_r$ . For an adapted Markov random field, the “deformed height function” corresponds to a point in the support as well, which will be the key to proving the required result.

We first set up some technical details. For  $x \in X_r$  and a finite set  $A \subset \mathbb{Z}^d$  denote:

$$\text{Range}_A(x) = \max_{n,m \in A} |\hat{x}_n - \hat{x}_m|,$$

where  $\hat{x} \in Ht$  is a height function associated with  $x$  by  $\phi(\hat{x}) = x$ .

**Lemma 7.2.** (*“Extremal values of height obtained on the boundary”*) *Let  $F \subset \mathbb{Z}^d$ , and suppose  $x \in X_r$  such that  $\text{Range}_{\partial F}(x) > 2$ . Then there exists  $y \in X_r$  such that  $y = x$  on  $F^c$  and*

$$\text{Range}_F(y) = \text{Range}_{\partial F}(y) - 2 = \text{Range}_{\partial F}(x) - 2.$$

*Proof.* Fix a height function  $\hat{x}$  such that  $\phi(\hat{x}) = x$ . Denote  $T := \max\{x_n : n \in \partial F\}$ ,  $B := \min\{x_n : n \in \partial F\}$ . Let  $k$  denote the sum of the absolute value for the deviations  $x_F$  from the interval  $(B, T)$ . That is

$$k := \sum_{n \in F} \max(x_n - T + 1, B - x_n + 1, 0),$$

We prove the claim by induction on this number  $k$ . If  $k = 0$  then  $y = x$  already satisfies the conclusion of this lemma because

$$\begin{aligned} \text{Range}_F(x) &= \max_{m \in F} \hat{x}_m - \min_{m \in F} \hat{x}_m \leq (\max_{m \in \partial F} \hat{x}_m - 1) - (\min_{m \in \partial F} \hat{x}_m + 1) \\ &= \text{Range}_{\partial F}(x) - 2. \end{aligned}$$

Now suppose  $k > 0$ , and let  $n \in F$  be a point where  $\hat{x}$  obtains an extremal value for  $F \cup \partial F$ . Without loss of generality suppose,

$$\hat{x}_n = \max_{m \in F \cup \partial F} \hat{x}_m.$$

Since all neighbors of  $n$  are in  $F \cup \partial F$ , it follows that  $\hat{x}_m = \hat{x}_n - 1$  for all  $m$  adjacent to  $n$ . Therefore the function  $\hat{y} \in \mathbb{Z}^d$  given by

$$\hat{y}_m = \begin{cases} \hat{x}_m - 2 & \text{for } m = n \\ \hat{x}_m & \text{otherwise} \end{cases}$$

is a valid height function. Hence we have *lowered* the height function at site  $n$ . Since  $\text{Range}_{\partial F}(x) > 2$ , it follows that  $\hat{y}_n$  is neither a minimum or a maximum for  $\hat{y}$  in  $F \cup \partial F$ . Thus, we can apply the induction hypothesis on  $y = \rho(\hat{y})$  and conclude the proof.  $\square$

**Lemma 7.3. (“Flat extension of an admissible pattern”)** Suppose  $x \in X_r$  and  $N \in \mathbb{N}$ . Then there exists  $y \in X_r$  such that  $y = x$  on  $D_{N+1}$  and

$$\text{Range}_{\partial D_{N+k}}(y) = \text{Range}_{\partial D_N}(x) - 2k,$$

for all  $1 \leq k \leq \frac{\text{Range}_{\partial D_N}(x)}{2}$ .

*Proof.* We will prove the following statement by induction on  $M \in \mathbb{N}$ : For any  $N \in \mathbb{N}$  and any height function  $\hat{x}$  on  $D_{N+1+M}$  with  $\text{Range}_{\partial D_N} \hat{x} = 2M$  there exist a height function  $\hat{y}$  on  $D_{N+1+M}$  such that  $\hat{y} = \hat{x}$  on  $D_{N+1}$  and for any  $1 \leq k \leq M$  and  $\text{Range}_{\partial D_{N+k}}(y) = 2M - 2k$ . Observe that the height function  $\hat{y}$  satisfies in particular  $\text{Range}_{\partial D_{N+M}}(y) = 0$ . Thus, the outermost boundary of  $\hat{y}$  is flat and it can be extended to a height function on  $\mathbb{Z}^d$ , so the lemma will follow immediately once we prove the statement above for all  $M \in \mathbb{N}$ .

For basis case of the induction, there is nothing to prove. Assume the result for some  $M \in \mathbb{N}$ . Let  $\hat{x} \in X_r$  be a height function  $\hat{y}$  on  $D_{N+1+(M+1)}$  such that  $\text{Range}_{\partial D_N} \hat{x} = 2(M+1)$ . Denote  $\tilde{N} := N + 1 + (M + 1) = N + M + 2$ . Let  $n \in D_{\tilde{N}} \setminus D_{N+1}$  be a site where  $\hat{x}$  obtains an extremal value for  $D_{\tilde{N}} \setminus D_N$ . If there is no such site then

$$\begin{aligned} \text{Range}_{\partial D_{N+1}}(\hat{x}) &= \max_{m \in \partial D_{N+1}} \hat{x}_m - \min_{m \in \partial D_{N+1}} \hat{x}_m \\ &= \left( \max_{m \in \partial D_N} \hat{x}_m - 1 \right) - \left( \min_{m \in \partial D_N} \hat{x}_m + 1 \right) \\ &= 2M \end{aligned}$$

proving the induction step for that case. Without the loss of generality we can assume that it is a maximum, that is,

$$\hat{x}_n = \max_{m \in D_{\tilde{N}} \setminus D_N} \hat{x}_m.$$

Then the function  $\hat{y}$  given by

$$\hat{y}_m = \begin{cases} \hat{x}_m - 2 & \text{if } m = n \\ \hat{x}_m & \text{otherwise} \end{cases}$$

is a valid height function on  $D_{\tilde{N}}$ . Hence we have lowered the height function at the site  $n$  of  $\hat{x}$ . Repeating the steps for sites with extremal height (formally, this is another internal induction), a height function  $\hat{z}$  can be obtained on  $D_{\tilde{N}}$  such that  $\hat{z} = \hat{x}$  on  $D_{N+1}$  and

$$\text{Range}_{\partial D_{N+1}}(\hat{z}) = 2M.$$

Since  $\tilde{N} = (N + 1) + 1 + M$ , we can apply the induction hypothesis to  $\hat{z}$ , substituting  $N + 1$  for  $N$ . Thus, we can now obtain a height function  $\hat{y}$  on  $D_{\tilde{N}}$  such that  $\hat{y} = \hat{x}$  on  $D_{N+1}$  and

$$\text{Range}_{\partial D_{N+k}}(\hat{y}) = \text{Range}_{\partial D_N}(\hat{y}) - 2k$$

for  $1 \leq k \leq \frac{\text{Range}_{\partial D_N}(\hat{y})}{2}$ .

This completes the proof of the statement.  $\square$

**Lemma 7.4. (“Patching an arbitrary finite configuration inside a non-steep point”)** Let  $r \geq 3$  be some integer and  $N, k \in \mathbb{N}$ . Choose  $y \in X_r$  which satisfies  $\text{Range}_{\partial D_{2N+2r+k+1}} y \leq 2k$ , and any  $x \in X_r$ . Then:

(1) If either  $r$  is odd or  $x_n - y_n$  is even for all  $n \in \mathbb{Z}^d$ , then there exists  $z \in X_r$  such that

$$z_n = \begin{cases} x_n & \text{if } n \in D_N \\ y_n & \text{if } n \in D_{2N+2r+k+1}^c \end{cases}$$

(2) Otherwise  $r$  is even and  $x_n - y_n$  is odd for all  $n \in \mathbb{Z}^d$ , then there exists  $z \in X_r$  such that

$$z_n = \begin{cases} x_{n+e_1} & \text{if } n \in D_N \\ y_n & \text{if } n \in D_{2N+2r+k+1}^c \end{cases}$$

The idea of this proof lies in the use of lemmata 7.2 and 7.3. Given any configuration on  $D_N$  we can extend it to a configuration on  $D_{2N}$  with flat boundary which than can be extended to  $y$  a little further away provided the range of the height function is not too large. A slightly different treatment is required for when  $r$  is even and when it is odd because the parity does not change according to the graph distance in the odd case.

*Proof.* By applying lemma 7.2  $k$  times we conclude that there exists  $y^{(1)} \in X_r$  such that  $y^{(1)} = y$  on  $D_{2N+2r+k+1}^c$  and

$$\text{Range}_{\partial D_{2N+2r+1}}(y^{(1)}) = 0.$$

Equivalently, there exists  $a \in \mathbb{Z}_r$  so that  $y_n^{(1)} = a$  for all  $n \in \partial D_{2N+2r+1}$ .

By lemma 7.3 choose  $x^{(1)} \in X_r$  such that  $x^{(1)} = x$  on  $D_N$  and

$$\text{Range}_{\partial D_{2N-1}}(x^{(1)}) = 0.$$

Equivalently, there exists  $b \in \mathbb{Z}_r$  so that  $x_n^{(1)} = b$  for all  $n \in \partial D_{2N-1}$ .

If we are in case (1), either  $r$  is even and  $b \equiv a \pmod{2}$  or  $r$  is odd. Either way, there is some integer  $k \in [0, \dots, r-1]$  such that  $a+k \equiv b-k \pmod{r}$ . Thus, we can find  $y^{(2)}, x^{(2)} \in X_r$ , so that  $y^{(2)}$  agrees with  $y^{(1)}$  outside  $D_{2N+2r}$ ,  $x^{(2)}$  agrees  $x^{(1)}$  in  $D_{2N}$ , and so that both  $x^{(2)}$  and  $y^{(2)}$  have a common constant value  $a+k = b+(r-k) \pmod{r}$  on  $\partial D_{2N+r-k}$ . Thus, we get the required  $z \in X_r$  by setting

$$z_n = \begin{cases} x_n^{(2)} & \text{for } n \in D_{2N+k} \\ y_n^{(2)} & \text{for } n \in D_{2N+k}^c \end{cases}$$

To prove case (2) we follow the same procedure, substituting  $x$  by  $\sigma_{e_1}x$ . □

We can now conclude the proof of proposition 7.1:

*Proof.* As in the proof of proposition 6.2, let  $\mu$  be a shift-invariant Markov random field and

$$v_j = \mu(\{x \in X_r : x_0 - x_{e_j} = 1 \pmod{r}\}) - \mu(\{x \in X_r : x_0 - x_{e_j} = -1 \pmod{r}\})$$

for  $j = 1, \dots, d$ . Assume that  $\text{supp}(\mu)$  is not frozen. Then by lemma 6.7,  $|v_j| < 1$  for all  $1 \leq j \leq d$ . Again, choose  $\epsilon > 0$  satisfying  $\epsilon < \frac{1}{4} \min\{1 - |v_j| : 1 \leq j \leq d\}$ .

We need to show that for any  $N \in \mathbb{N}$ , and any configuration  $c \in B_{D_N}(X_r)$ ,  $\mu([c]_{D_N}) > 0$ .

Using corollary 6.4 it follows that for sufficiently large  $k$ ,

$$\mu(\{y \in X_r : \text{Range}_{\partial D_k}(y) \leq 2(1 - \epsilon)k\}) > 1 - \epsilon.$$



Now choose  $k > (2N + 2r + 1)$  large enough so that for some  $y \in X_r$  with

$$\text{Range}_{\partial D_k}(y) \leq 2(1 - \epsilon)k \leq 2(k - (2N + 2r + 1))$$

and  $\mu([y]_{D_{k+1}}) > 0$ . By lemma 7.4, it follows that there exists  $z \in X_r$  with  $z_n = y_n$  for  $n \in \mathbb{Z}^d \setminus D_k$  and  $z_n = c_n$  for  $n \in D_N$ . Since  $\mu$  is an adapted Markov random field, it follows that  $\mu([z]_{D_{k+1}}) > 0$ , and in particular  $\mu([c]_{D_N}) > 0$ . □

## 8. FULLY SUPPORTED SHIFT-INVARIANT GIBBS MEASURES ON $X_r$

Next we demonstrate the existence of a fully-supported shift-invariant Gibbs measure for shift invariant nearest neighbor interactions on  $X_r$ . By proposition 7.1, this is equivalent to the existence of a non-frozen Gibbs measure. We will obtain these by showing that equilibrium measures for certain interactions are non-frozen. To state and prove this result, we need to introduce the notion of (measure-theoretic) pressure and equilibrium measures and apply a theorem of Lanford and Ruelle relating equilibrium measures and Gibbs measures. Our presentation is far from comprehensive, and is aimed to bring only definitions necessary for our current results. We refer readers seeking background on pressure and equilibrium measures to the many existing textbooks on the subject, for instance [19, 25].

Let  $\mu$  be a shift-invariant probability measure on a shift of finite type  $X$ . The *measure theoretic entropy* can be defined by

$$h_\mu = \lim_{i \rightarrow \infty} \frac{1}{|D_i|} H_\mu^{D_i},$$

where  $H_\mu^{D_i}$  is the Shannon-entropy of  $\mu$  with respect to the partition of  $X$  generated by  $D_i$ , the definition of which is given by:

$$H_\mu^{D_i} = \sum_{a \in \mathcal{B}_{D_i}} -\mu([a]_{D_i}) \log \mu([a]_{D_i}),$$

with the understanding that  $0 \log 0 = 0$ .

Given a continuous function  $f : X \rightarrow \mathbb{R}$ , the *measure-theoretic pressure* of  $f$  with respect to  $\mu$  is given by

$$P_\mu(f) = \int f d\mu + h_\mu.$$

A shift-invariant probability measure  $\mu$  is an *equilibrium state* for  $f$  if the maximum of  $\nu \mapsto P_\nu(f)$  over all shift-invariant probability measures is obtained at  $\mu$ . The existence of an equilibrium state for any continuous  $f$  follows from upper-semi-continuity of the function  $\nu \mapsto P_\nu(f)$  with respect to the weak-\* topology.

Let  $\phi$  be a nearest neighbor interaction on  $X$ . Define a function  $f_\phi : X \rightarrow \mathbb{R}$  by

$$f_\phi(x) = \sum_{\substack{A \text{ finite} \\ |(0,0) \in A \subset \mathbb{Z}^d|}} \frac{1}{|A|} \phi(x|_A).$$

The following is a restricted case of a classical theorem by Lanford and Ruelle:

**Theorem. (Lanford-Ruelle theorem [14])** *Let  $X$  be a  $\mathbb{Z}^d$ -shift of finite type and  $\phi$  a shift-invariant nearest neighbor interaction. Then any equilibrium state  $\mu$  for the  $f_\phi$  is a Gibbs state for the given interaction  $\phi$ .*

The *topological entropy* of a  $\mathbb{Z}^d$ -subshift  $X$  can be defined by

$$h(X) = \lim_{k \rightarrow \infty} \frac{1}{|D_k|} \log |\mathcal{B}_{D_k}(X)|.$$

We recall that the well known *variational principle* for topological entropy of  $\mathbb{Z}^d$ -actions, which (in particular) asserts that  $h(X) = \sup_{\nu} h_{\nu}$  whenever  $X$  be a  $\mathbb{Z}^d$ -shift space and the supremum is over all probability measures on  $X$ .

**Lemma 8.1.** *Let  $M$  be a Gibbs cocycle on  $X_r$  with a shift-invariant nearest neighbor interaction. Then there exists a shift-invariant nearest neighbor interaction  $\phi$  such that  $M = M_{\phi}$  and any equilibrium measure for  $f_{\phi}$  is non-frozen.*

*Proof.* Let  $(x^{(i)}, y^{(i)}) \in \Delta_{X_r}$  be as in the proof of proposition 5.2. If  $M \in \mathcal{G}_{X_r}$  then there exists a shift-invariant nearest neighbor interaction  $\phi$  so that

$$\begin{aligned} M(x^{(i)}, y^{(i)}) &= \phi([i+2]_0) - \phi([i]_0) \\ &+ \sum_{j=0}^{r-1} (\phi([i+2, i+1]_j) - \phi([i+1, i]_j) + \phi([i+1, i+2]_j) - \phi([i, i+1]_j)). \end{aligned}$$

Furthermore, we can choose an interaction  $\phi$  to be such that

$$\phi([i, i+1]_j) = \phi([i+1, i]_j) = a_i \text{ for all } i \in \mathbb{Z}_r \text{ and } j \in \{1, 2, \dots, d\}$$

and  $\phi([i]_0) = 0$  for all  $i \in \mathbb{Z}_r$ . For such an interaction we have

$$\begin{aligned} \int f_{\phi}(x) d\mu(x) &= \sum_{j=1}^d \sum_{i=0}^{r-1} \phi_j(i, i+1) \mu([i, i+1]_j) + \phi_j(i+1, i) \mu([i+1, i]_j) \\ &= \sum_{j=1}^d \sum_{i=0}^{r-1} a_i (\mu([i, i+1]_j) + \mu([i+1, i]_j)). \end{aligned}$$

Let  $a = \max_{1 \leq i \leq r} a_i$  attained by  $a_{i_0}$ . It follows that for any shift-invariant probability measure

$$\int f_{\phi}(x) d\mu(x) \leq d \cdot a$$

with equality holding iff  $\mu([i, i+1]_j) = \mu([i+1, i]_j) = 0$  for all  $a_i < a$  and  $j = 1, \dots, d$ .

For a frozen measure  $\mu$  it follows that for some  $j \in \{1, 2, \dots, d\}$ ,  $\mu([i, i+1]_j) > 0$  or  $\mu([i+1, i]_j) > 0$  for all  $i \in \{0, 1, \dots, r-1\}$ . Thus if  $a_i < a$  for *some*  $0 \leq i \leq r-1$ , it follows that for any frozen measure  $\mu$ ,

$$\int f_{\mu}(x) < \sup_{\nu} \int f d\nu(x).$$

where the supremum is attained by the measure supported on the orbit of the point periodic point  $x \in X_r$  given by

$$x_n = \begin{cases} i_0 & \|n\|_1 \text{ odd} \\ i_0 + 1 & \|n\|_1 \text{ even} \end{cases}$$

It is easily verified that if  $\mu$  is frozen then  $h_{\mu} = 0$ . Thus in this case, for any frozen probability measure  $\mu$

$$P_{\mu}(f_{\phi}) = \int f_{\phi}(x) \mu(x) < \sup_{\nu} \int f_{\phi}(X) d\nu(x) \leq \sup_{\nu} P_{\nu}(f_{\phi})$$

and in particular any frozen measure  $\mu$  can not be an equilibrium measure for  $f_V$ .

The remaining case is when  $a_i = a$  for all  $i$ , in which case  $f_\phi(x) = d \cdot a$  is constant. Thus, by the variational principle  $\sup_\nu P_\nu(f) = d \cdot a + \sup_\nu h_\nu = d \cdot a + h(X_r)$ . Since  $h(X_r) > 0$ , it follows that the strict inequality  $P_\mu(f_\phi) < \sup_\nu P_\nu(f_\phi)$  holds also in this case for any frozen measure  $\mu$ . Thus we have the result that for a given Gibbs cocycle with a shift-invariant nearest neighbor interaction there exists an interaction for that cocycle such that the corresponding equilibrium state is not frozen. By proposition 7.1 the proof is complete.  $\square$

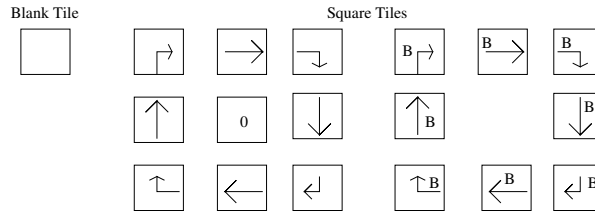
**Corollary 8.2.** *For any shift-invariant Gibbs cocycle  $M$  on  $X_r$  there exists a shift-invariant nearest neighbor interaction  $\phi$  on  $X_r$  with  $M = M_\phi$ , and a corresponding shift-invariant Gibbs state  $\nu$  with  $\text{supp}(\nu) = X_r$ .*

*Proof.* By lemma 8.1, there exists a nearest neighbor interaction  $V$  on  $X_r$  with  $M = M_\phi$  and an equilibrium measure  $\mu$  for  $f_\phi$  which is non-frozen. By the Lanford-Ruelle theorem such  $\mu$  is a Gibbs state for  $\phi$ .  $\square$

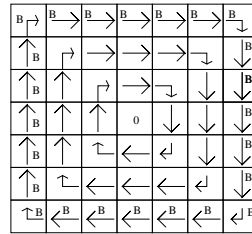
## 9. “STRONGLY” NON-GIBBSIAN SHIFT-INVARIANT MARKOV RANDOM FIELDS

In this section we demonstrate the existence of shift-invariant Markov random fields for whose specification is not given by any shift-invariant finite range interaction. Our example proves that in general, in contrast to shift-invariant finite range interactions, the specification of a Markov random field cannot be “given by a finite number of parameters”.

Let  $\mathcal{A}$  be the alphabet consisting of the following 18 tiles, shown in the figure below: There is a blank tile, a “seed” tile (marked in the figure with “0”), 8 “interior arrow tiles”, 4 of which correspond to the cardinal directions and 4 corresponding to “corner directions” and 8 additional “border arrow tiles” (marked in the figure with and extra mark “B”).

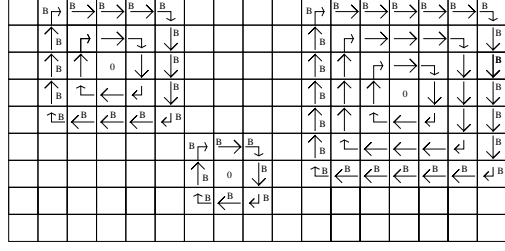


All tiles other than the blank tile will be called *square tiles* and the tiles with arrows will be called *arrow tiles*. The arrow tiles with ‘B’ will be called *border tiles* and those without ‘B’ will be called *interior tiles*. Configurations of an  $(2n + 1) \times (2n + 1)$  square whose inner boundary consists of border tiles as shown in the figure below will be called an *n-square*:



,that is configurations surrounded by border tiles will be referred to as *squares*. The idea is to have the interior tiles form squares with a seed tile in the center and border tiles on the

boundary “floating in the sea of the blank tiles”. A configuration in general should look like the figure below.



Let  $X$  be the nearest neighbor shift of finite type on alphabet  $\mathcal{A}$  with constraints given by

- (1) Any “arrow head” must meet an “arrow tail” of matching type.
- (2) Adjacent arrow tiles should not point in opposite directions.
- (3) Two corner direction tiles cannot be adjacent to one another.
- (4) The seed tile is surrounded by arrow tiles on all sides.
- (5) An interior tile is always surrounded by other square tiles while a border tile has an interior tile on its right and the blank tile on its left.

Clearly  $X$  has positive entropy. Denote by  $B_n \subset \mathbb{Z}^2$  the set  $\{(i, j) \in \mathbb{Z}^2 \mid |i|, |j| \leq n\}$ .

**Proposition 9.1.** *Let  $\mu$  be any measure of maximal entropy for  $X$ . Then  $\mu$  is fully-supported.*

*Proof.* Let  $\mu$  be a measure of maximal entropy for  $X$ . Our first observation is that  $\mu$ -almost surely there is no infinite connected component composed of square tiles. An infinite connected component can be either an infinite sized square, a corner or a periodic point composed of one of the arrow tiles. Being transient events, by Poincare recurrence theorem the probability of having a corner or an infinite sized square is 0. Since  $\mu$  is a measure of maximal entropy for  $X$ , the measure of periodic points is 0 as well. Thus the observation has been verified. Let  $x \in X$  be an element such that it does not have any infinite connected component composed of square tiles. We will now show for all  $r \in \mathbb{N}$  that there exists a finite set  $A_r$  such that  $B_r \subset A_r \subset \mathbb{Z}^2$  and  $x_i$  is the blank tile for all  $i \in \partial A_r$ . Let  $Sq_1 \in \mathcal{A}^{C_1}, Sq_2 \in \mathcal{A}^{C_2}, \dots, Sq_k \in \mathcal{A}^{C_k}$  be an enumeration of the squares in  $x$  such that  $C_i \cap B_{r+1} \neq \emptyset$ . Let

$$A_r = \cup_{i=1}^k C_i \cup B_r.$$

Since every square is surrounded by the blank tile,  $A_r$  has the required properties.

Now consider any element  $y \in X$ . By completing the incomplete square in  $y|_{B_n}$  we obtain a  $z \in X$  such that

$$z_i = \begin{cases} y_i & \text{for } i \in B_n \\ \text{blank tile} & \text{for } i \in B_{4n}^c. \end{cases}$$

Now choose any  $x \in \text{supp}(\mu)$  which does not have any infinite connected component composed of square tiles. By the previous argument, find  $A_{4r} \subset \mathbb{Z}^2$  such that  $B_{4r} \subset A_{4r}$  and  $x_i$  is the blank tile for all  $i \in \partial A_{4r}$ . Then  $z = x$  on  $\partial A_{4r}$ . By Lanford-Ruelle theorem  $\mu$  is a Markov random field with the uniform specification adapted to  $X$ . Therefore

$$\frac{\mu([z]_{A \cup \partial A})}{\mu([x]_{A \cup \partial A})} = 1$$

proving  $\mu([z]_{B_n}) = \mu([y]_{B_n}) > 0$ . Since this is true for any  $n \in \mathbb{N}$  we can conclude  $y \in \text{supp}(\mu)$ . But the choice of  $y \in X$  is arbitrary. Hence  $\text{supp}(\mu) = X$ .  $\square$

To obtain a space such the dimension of the Markov cocycles is infinite let  $Y$  be a nearest neighbor shift of finite type with the alphabet as in figure 1 but now with two types of square tiles, type 1 and 2. The adjacency rules are as in the subshift  $X$  but also that adjacent square tiles must be of the same type, that is, any square in an element of  $Y$  will consist entirely of tiles of type 1 or of type 2. Let  $\mathbf{p} \in (0, 1)^\mathbb{N}$  and  $\phi : Y \rightarrow X$  be the map which forgets the type of square tiles. We will now construct a shift-invariant Markov random field  $\mu_{\mathbf{p}}$  obtained by picking a particular configuration in  $X$  according to a measure of maximal entropy  $\mu$  and then choosing the type of every  $i$ -square to be 1 with probability  $p_i$  and 2 with probability  $1 - p_i$ . A more precise description follows. Let  $\mathcal{F} = \phi^{-1}(\mathcal{B}(X))$  be the pull-back of the Borel sigma-algebra on  $X$ . For any  $y \in Y$ ,  $i \in \mathbb{N}$  and  $\Lambda \subset \mathbb{Z}^2$  finite consider the functions

$$\begin{aligned} m_\Lambda^i(y) &= \text{the number of } i\text{-squares of type 1 in } y \text{ intersecting } \Lambda \\ \text{and} \\ n_\Lambda^i(y) &= \text{the number of } i\text{-squares of type 2 in } y \text{ intersecting } \Lambda. \end{aligned}$$

$\mu_{\mathbf{p}}$  is the unique probability measure on  $Y$  satisfying  $\mu_{\mathbf{p}}|_{\mathcal{F}} = \mu \circ \phi^{-1}$  and

$$\mu_{\mathbf{p}}([y]_\Lambda | \mathcal{F}) = \prod_{i=1}^{\infty} p_i^{m_\Lambda^i(y)} (1 - p_i)^{n_\Lambda^i(y)}$$

for all  $\Lambda \subset \mathbb{Z}^2$  finite,  $y \in Y$ .

Note that  $\mu$ -almost surely configurations in  $X$  has no infinite-sized squares so  $\tilde{\mu}$ -almost surely the  $m_i$ 's and  $n_i$ 's account for all the squares intersecting  $\Lambda$ .

**Proposition 9.2.** *For any  $p \in (0, 1)^\mathbb{N}$ , the measure  $\mu_{\mathbf{p}}$  defined above is a shift-invariant Markov random field.*

*Proof.* Let  $(y, y') \in \Delta_Y$  and  $F = \{i \in \mathbb{Z}^2 \mid y_i \neq y'_i\}$ . Then

$$\begin{aligned} m_{F \cup \partial F}^i(y) - m_{F \cup \partial F}^i(y') &= m_\Lambda^i(y) - m_\Lambda^i(y') \\ n_{F \cup \partial F}^i(y) - n_{F \cup \partial F}^i(y') &= n_\Lambda^i(y) - n_\Lambda^i(y'). \end{aligned}$$

for  $\Lambda \supset F \cup \partial F$  finite. Therefore it is sufficient to prove that  $M_{\mathbf{p}} : \Delta_Y \rightarrow \mathbb{R}$  given by

$$\begin{aligned} M_{\mathbf{p}}(y, y') &= \log \frac{\mu_{\mathbf{p}}([y']_{F \cup \partial F} | \mathcal{F})(y')}{\mu_{\mathbf{p}}([y]_{F \cup \partial F} | \mathcal{F})(y)} \\ &= \sum_{i=1}^{\infty} (m_{F \cup \partial F}^i(y') - m_{F \cup \partial F}^i(y)) \log(p_i) + (n_{F \cup \partial F}^i(y') - n_{F \cup \partial F}^i(y)) \log(1 - p_i) \end{aligned}$$

for all  $(y, y') \in \Delta_Y$  and  $F = \{i \in \mathbb{Z}^2 \mid y_i \neq y'_i\}$  defines a shift-invariant Markov cocycle. Let  $i \in \mathbb{N}$  be chosen. We will verify that  $M : \Delta_Y \rightarrow \mathbb{R}$

$$M(y, y') = m_{F \cup \partial F}^i(y) - m_{F \cup \partial F}^i(y')$$

for  $F, y, y'$  as above defines a shift-invariant Markov cocycle. By linearity of shift-invariant Markov cocycles this is sufficient.

Given  $(y, y'), (y', y'') \in \Delta_Y$ ,

$$F = \{i \in \mathbb{Z}^2 \mid y_i \neq y'_i \text{ or } y'_i \neq y''_i\}$$

and  $\Lambda \supset F \cup \partial F$  finite it follows that

$$M(y, y') + M(y', y'') = m_\Lambda^i(y) - m_\Lambda^i(y') + m_\Lambda^i(y') - m_\Lambda^i(y'') = M(y, y'').$$

This proves that  $M$  is a shift-invariant cocycle. To complete the proof we need to verify that for any  $(y, y') \in \Delta_Y$ ,  $M(y, y')$  depends only on the sites where  $y$  and  $y'$  differ and its boundary. As before let  $(y, y') \in \Delta_Y$  and  $F = \{i \in \mathbb{Z}^2 \mid y_i \neq y'_i\}$ . For each square in  $y$  intersecting  $F \cup \partial F$  there is either a crossing between the horizontal sides in  $F \cup \partial F$  or a crossing between the vertical sides in  $F^c$ . Here a crossing is a path from one side of the square to its opposite side where diagonal steps are allowed. In the former case the subconfiguration on  $F \cup \partial F$  determines the entire square and in the latter the subconfiguration of  $F^c$  does the same. Consider the function  $g_i^\Lambda$  such that  $g_i^\Lambda(x)$  is the number of  $i$ -squares of type 1 intersecting  $\Lambda$  in  $x$  which are determined by the subconfiguration on  $\Lambda$  for every  $x \in Y$ . Then it follows that

$$M(y, y') = m_{F \cup \partial F}^i(y) - m_{F \cup \partial F}^i(y') = g_{F \cup \partial F}^i(y) - g_{F \cup \partial F}^i(y')$$

is dependent exactly on  $y|_{F \cup \partial F}$  and  $y'|_{F \cup \partial F}$ .  $\square$

This proves existence of uncountably many linearly independent shift-invariant Markov cocycles which have corresponding fully-supported shift-invariant Markov random fields. Since the space of Markov cocycles which come from some finite range shift-invariant is a union of finite dimensional vector spaces, this further implies that there exists a shift-invariant Markov random field which is not Gibbs for any shift-invariant finite interaction. Alternatively, note that for any Gibbs cocycle with shift-invariant finite range interaction the magnitude of the cocycle at a particular homoclinic pair is at most linear in the size of the sites at which the two configurations differ. We can choose a  $\mathbf{p} \in (0, 1)^\mathbb{N}$  such that this does not happen.

A simple variation on the above construction yields topological Markov fields which are not sofic: Choose  $\mathbf{p} \in [0, 1]^\mathbb{N}$ . If  $p_i = 0$  or  $p_i = 1$ , this would disallow squares of certain types for specific sizes. Each such  $\mathbf{p}$  would determine a shift-invariant Markov random field supported on a shift space contained in  $Y$ . Since there are uncountable many such subshifts many such spaces would be not sofic. However it is easy to see from the proofs above that these are global topological Markov fields.

## 10. CONCLUSION AND FURTHER PROBLEMS

In this paper we demonstrated the applicability of Markov cocycles to studying Markov random fields and Gibbs states with nearest neighbor interaction. In cases where the space of shift-invariant Markov cocycles is finite dimensional, these can act as a substitute for nearest neighbor interactions in providing a description of the specification “using finitely many parameters”. There are several questions which remain to be answered. We mention a few here.

**Question 1:** What characterizes the property of being the support of a shift-invariant Markov random field on the Calyey graph of  $\mathbb{Z}^d$ ? We know it is equivalent to being a non-wandering nearest neighbor shift of finite type when  $d = 1$ . For finite graphs, it suffices to be a topological Markov field. In higher dimensions we wonder whether it is sufficient to have

a topological Markov field with a shift-invariant probability measure supported on it. Also: Suppose a  $\mathbb{Z}^d$  shift of finite type is the support of some shift-invariant Markov random field. Must it also be the support of a shift-invariant Gibbs measure for some nearest neighbor interaction?

**Question 2:** In this paper we prove that the dimension of the space of shift-invariant Markov cocycles on  $X_r$  is  $r$  and that of Gibbs cocycles with shift-invariant nearest neighbour interactions is  $r - 1$ . Moreover we prove that any shift-invariant Markov random field adapted to these spaces has a shift-invariant nearest neighbour interaction. Finally we prove that corresponding a given shift-invariant Markov cocycle there exists a fully supported shift-invariant Markov random field if and only if it is Gibbs with a shift-invariant nearest neighbour interaction. Suppose we are given a nearest neighbor shift of finite type with the pivot property along with its globally allowed patterns on  $\{0\} \cup \partial\{0\}$ . Is there an algorithm to determine the dimension of the space of shift-invariant Markov cocycles? If so, is there a way to decide which of these cocycles have a corresponding fully supported shift-invariant probability measure on the subshift? In case the subshift has a safe symbol, such an algorithm can be derived via from the proof of the Hammersley-Clifford theorem [7] and also from Lemma 3.1 in [8]. Note that it is not hard to see that given this information there is an algorithm to determine the dimension of the space of Gibbs cocycles with shift-invariant finite range interactions. Specific models of interest would be  $r$ -colorings of  $\mathbb{Z}^2$  with  $r \geq 6$  and domino tilings of  $\mathbb{Z}^2$  (with the generalized pivot property). One can also question how much does the underlying graph effect these results. For instance, we know that that every shift-invariant Markov random field is Gibbs with a shift-invariant nearest neighbour interaction when the underlying graph is  $\mathbb{Z}$  [6] and every Markov random field is Gibbs with a nearest neighbour interaction when the underlying graph is finite and decomposable, that is, a graph which can be decomposed into triangles and edges[15].

**Question 3:** In section 9, we construct a subshift such that the dimension of the space of shift-invariant Markov cocycles is uncountable. However this space does not have any nice mixing property like block gluing or strongly irreducible[2]. Does there exists a subshift which is strongly irreducible or even block gluing such that the dimension of the space of Markov cocycles is uncountable?

#### ACKNOWLEDGEMENTS

We will like to thank Prof. Brian Marcus, Prof. Ronnie Pavlov and Prof. Mike Boyle for many a useful discussions. The first author is extremely grateful to Prof. Brian Marcus, his PhD advisor at the University of British Columbia under whose tutelage he has learnt all the ergodic theory that he currently knows. He will also like to thank the University of British Columbia for providing the necessary funding and opportunity to conduct this research. The second author would like to thank Prof. Brian Marcus, PIMS and the University of British Columbia for hosting him as a post-doctoral fellow during which period the problems in this paper were studied.

#### REFERENCES

- [1] M. B. Avez. Description of Markov random fields by means of Gibbs conditional probabilities. *Teor. Veroyatnost. i Primenen.*, 17:21–35, 1972.
- [2] Mike Boyle, Ronnie Pavlov, and Michael Schraudner. Multidimensional sofic shifts without separation and their factors. *Transactions of the American Mathematical Society*, 362(9):4617–4653, 2010.

- [3] Graham R Brighwell and Peter Winkler. Gibbs measures and dismantable graphs. *Journal of Combinatorial Theory, Series B*, 78(1):141–166, 2000.
- [4] Robert Burton and Jeffrey E Steif. Non-uniqueness of measures of maximal entropy for subshifts of finite type. *Ergodic Theory Dynam. Systems*, 14(02):213–235, 1994.
- [5] Nishant Chandgotia. Markov random fields and measures with nearest neighbour potentials. *MSc Thesis*, 2011.
- [6] Nishant Chandgotia, Guangyue Han, Brian Marcus, Tom Meyerovitch, and Ronnie Pavlov. One dimensional markov random fields, markov chains and topological markov fields. *AMS Proceedings(To Appear)*, 2011.
- [7] C.J.Preston. *Gibbs states on countable sets*. Number 68 in Cambridge Tracts in Mathematics. Cambridge University Press, 1974.
- [8] Serguei Dachian and BS Nahapetian. Description of random fields by means of one-point conditional distributions and some applications. In *Markov Proc. Rel. Fields*, volume 7, pages 193–214, 2001.
- [9] R.L. Dobruschin. Description of a random field by means of conditional probabilities and conditions for its regularity. *Theor., Prob. Appl.*, 13:197–224, 1968.
- [10] Noam Elkies, Greg Kuperberg, Michael Larsen, and James Propp. Alternating-sign matrices and domino tilings (part i). *Journal of Algebraic Combinatorics*, 1(2):111–132, 1992.
- [11] H. Georgii. *Gibbs measures and phase transitions*. de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, 1988.
- [12] Dan Gieger, Christopher Meek, and Bernd Sturmfels. On the toric algebra of graphical models. *Ann. Statist.*, 34(3):1461–1492, 2006.
- [13] J. M. Hammersley and P. Clifford. Markov field on finite graphs and lattices. 1971.
- [14] O.E. Lanford and D. Ruelle. Observables at infinity ad states with short range correlation in statistical mechanics. *Commun. Math. Phys.*, 13:194–215, 1969.
- [15] S. Lauritzen. *Graphical Models*. Clarendon Press, Oxford, 1996.
- [16] J. Moussouris. Gibbs and markov random systems with constraints. *Journal of Statistical Physics*, 10:11–33, 1974.
- [17] Lars Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev. (2)*, 65:117–149, 1944.
- [18] Karl Petersen and Klaus Schmidt. Symmetric Gibbs measures. *Trans. Amer. Math. Soc.*, 349(7):2775–2811, 1997.
- [19] D. Ruelle. *Thermodynamic formalism*. Addison-Wesley, 1978.
- [20] Klaus Schmidt. The cohomology of higher-dimensional shifts of finite type. *Pacific J. Math.*, 170(1):237–269, 1995.
- [21] Klaus Schmidt. Invariant cocycles, random tilings and the super- $K$  and strong Markov properties. *Trans. Amer. Math. Soc.*, 349(7):2813–2825, 1997.
- [22] Klaus Schmidt. Tilings, fundamental cocycles and fundamental groups of symbolic  $\mathbf{Z}^d$ -actions. *Ergodic Theory Dynam. Systems*, 18(6):1473–1525, 1998.
- [23] S. Sherman. Markov random fields and Gibbs random fields. *Israel J. Math.*, 14:92–103, 1973.
- [24] Frank Spitzer. Markov random fields and gibbs ensembles. *Amer. Math. Monthly*, 78:142–154, 1971.
- [25] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, 1982.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA,CANADA  
*E-mail address*: nishant@math.ubc.ca

DEPARTMENT OF MATHEMATICS, BEN-GURION UNIVERSITY OF THE NEGEV, ISRAEL  
*E-mail address*: mtom@math.bgu.ac.il